ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

Library Trends

Issues and Problems in Designing a National Program of Library Automation

HENRY J. DUBESTER
Issue Editor

April, 1970

Library Trends

A Publication of the University of Illinois Graduate School of Library Science

LIBRARY TRENDS, a quarterly journal of librarianship, provides a medium for evaluative recapitulation of current thought and practice, searching for those ideas and procedures which hold the greatest potentialities for the future.

Each issue is concerned with one aspect of librarianship. Each is planned with the assistance of an invited advisory editor. All articles are by invitation. Suggestions for future issues are welcomed and should be sent to the Managing Editor.

Published four times a year, in July, October, January, and April. Office of Publication: University of Illinois Graduate School of Library Science, Urbana, Illinois. Entered as second-class matter June 25, 1952, at the Post Office at Urbana, Illinois, under the act of August 24, 1912.

Subscription price is $8.00 a year. Individual issues are priced at $2.50. All foreign subscriptions and orders should be accompanied by payment. Address orders to Subscription Department, University of Illinois Press, Urbana, Illinois. Editorial correspondence should be sent to LIBRARY TRENDS, University of Illinois Graduate School of Library Science, Urbana, Illinois. Indexed in *Library Literature, Library Science Abstracts,* and *PAIS.*

Library Trends

VOLUME 18 • NUMBER 4

APRIL, 1970

Issues and Problems in Designing a National Program of Library Automation

HENRY J. DUBESTER
Issue Editor

CONTRIBUTORS TO THIS ISSUE

This Page Intentionally Left Blank

Introduction

HENRY J. DUBESTER

IN THE PAGES that follow, some key issues and problems are explored and explicated with respect to the possibilities and requirements of a national program looking toward library automation. Library automation is a subject that has a good grip on the imagination and energies of an increasing number of librarians. Nevertheless, it is a young effort even as these are counted in this rapidly growing and changing field. Very few individuals within the library community can address the problems posed in the title of this issue of *Library Trends* with either confidence or expertise. The authors of the articles that follow can be numbered among such experts. It seems appropriate, however, to describe some of the steps that were taken to enlist their cooperation.

The initial overture by the editor of this issue of *Library Trends* was cast in terms of a prospective title describing an issue devoted to "System Design for a National Program of Library Automation." That title was rejected in favor of the present one which doubtless proved more felicitous in several ways. It removed concern that the collection of articles would yield a system design. It also shifted the focus to problems and issues that would need resolution no matter what the design of some eventual program for library automation might be. This tended to free prospective authors from time constraints and to allow them to view the problem set in a broader context.

What are the salient problems and significant issues which would retain essential validity irrespective of the specific character of a national program for library automation? It seemed reasonable to identify these problems and issues before seeking authors to write about them.

The advice of the library community, through opinions solicited from individuals who could address this question, proved invaluable. Opinions were sought and received from Scott Adams, Deputy Director of

Henry J. Dubester is Deputy Head, Office of Science Information Services, National Science Foundation, Washington, D.C.

the National Library of Medicine; Verner W. Clapp, Consultant to the Council on Library Resources, Inc.; Richard De Gennaro, Associate University Librarian for System Development at Harvard University; Herman H. Fussler, Director of the University of Chicago Library; L. Quincy Mumford, Librarian of Congress; James E. Skipper, while still Associate University Librarian at Princeton University; Vladimir Slamecka, Director of the School of Information Science at the Georgia Institute of Technology; Charles H. Stevens, a leading staff member on Project Intrex at the Massachusetts Institute of Technology; and Melvin J. Voigt, University Librarian at the University of California, San Diego. The responses reflected the keen interest of the above-named "librarians" and also validated the usefulness of pursuing the task of collecting writings on this topic. There was considerable overlap in the expressions and judgments reflected in the assembled replies. Their assessments, however, yielded an outline which was accepted by the Publications Board of the University of Illinois School of Library Science and which was also provided to each invited author in order to describe a framework of reference within which his contribution would fit. The outline of topics was rather definite at this point in time; the gloss associated with each could be only conjectural and would inevitably be modified by the author. It should be useful to display this initial framework.

Introduction—Where are we with respect to a national program of library automation? Where are we going and what are the necessary milestones toward the development of such a program? Essentially, this article will provide an assessment of the present situation, of ongoing activities, and an appreciation of present outlook.

Standards—There is outstanding agreement on the requirement for standards in any future automated library system and therefore agreement on their necessity as part of the program leading to such a system. The explication of standardization requirements in library activities pursued in a computerized mode will be the focus of this article.

Economics—The costs of automation played against the resources of libraries is one aspect of this topic. Another aspect is the economic value of automated libraries. The latter will be difficult to treat. The economic resources are variable when played against different types of libraries. Cost/benefit relationships are at issue here.

Manpower—The requirements of the library community, faced with an automation program on a national scale, for manpower trained in

operating, developing, and doing research in computerized information processing poses a training and manpower allocation problem which is the subject of growing concern and warrants consideration.

Networks—There is much current concern with impending networks as a consequence of automation and computerized transfer of information and data. Everyone speaks of networks but no one knows what they are. There are questions of geographic, functional and other relationships as well as issues concerning roles and responsibilities with respect to network operations and organization, and the topic of initial development. In one sense, the entire issue of *Library Trends* revolves around this topic. The problem will be to isolate the network aspects from the other topics listed.

Hardware—Computer technology and its hardware will play a continuing decisive role in any developing program of automation. What role does computer hardware play? Is the technology ahead of the requirements that the library community is capable of explicating? What are the inherent promises of computers to which libraries must accommodate in order to exploit the technology in a maximum manner?

Sectoral Relationships—The intent of this topical heading is to allow exploration of the role and responsibilities of government, academia, professional societies, libraries, etc., in relationship to a national program. Assuming that economic resources require intervention and support of the federal government and possibly that of state and local government bodies, and further assuming that resources will be contributed by the other sectors as well, what are the conditions that need to be met to harmonize the various protagonists, provide for program development, facilitate a decision process which at the same time insures continuing cooperation and initiative of all participants?

Retrospective Conversion—This topic was suggested by several of the correspondents. The issue is simple: an automated system will require that the entire bibliographic record exist in machine-readable form. Retrospective conversion of the existing record is necessary. How can this be accomplished? Centrally? Cooperatively? What about costs? Is the above assertion indisputable?

Research—The anticipated national program is viewed not simply as a one-shot affair. It will be a continuing effort which will need support and depend upon the results of research specifically focused on library automation. This research is presently fitful and scattered in its focus.

This was the outline. Each author had this overview. Each author

on request provided an abstract of the promised article, and each author received all the abstracts. As the papers were received, they also were forwarded to each author. These articles stand as conceived by their authors. The conformance or departure from the topical outline reflects their own interpretations and insights. Some generalized inferences can be made.

The interrelationships between problems or issues that are conceptually separable becomes ever more clearly evident as the individual articles demonstrate. Research and its relevance to library automation is a recurring theme as one peruses the treatment of "hardware," "economics," "networks," and other topics in addition to the article on research itself. The problem of standards becomes a recurring refrain, as does the issue of the locus of responsibility for the component aspects of developing a national plan and resolving the problems and issues posed by the articles. The inference could also be permitted that we, the library community, are still quite distant from an effective, coherent national program leading to the design of a program for library automation. But we are moving closer; indeed, have moved closer even within the brief period spanning the conception of this issue and its preparation for publication. There is no concerted effort which unites the library community with respect to all the problems subsumed under the term "standardization." Nevertheless, the progress of the MARC format as a means of exchanging machine-readable catalog information and as a national and even international standard is impressive. The economics of automation are not sufficiently understood to provide concrete foundations for valuation of library services in the context of cost/benefit analysis; the recognition of data requirements has progressed, however, and experience is beginning to accumulate to facilitate analysis. Manpower continues to be a vexing problem, but increasingly library schools are generating new curricula that confront traditional librarianship with a new breed of information scientists. Although library networks are not really extant, and network planning still lies ahead, the availability of bibliographic information in a form that can be processed by the computer is providing real pressure on libraries to face the requirement to exploit the investment in this capability. Here libraries are faced with the reality of competition from information centers that have begun to provide "on-line" information service based on digitalized data bases provided by indexing and abstracting services. Some libraries have taken initial steps to assimilate such services within their own operations. Similarly, hardware

developments are becoming less alien to the traditional library. Although few have their own computers, many are utilizing computers for portions of the internal operations. This, in turn, affects the development and training of library manpower and has significantly reduced the scare value of the computer as a replacement for the librarian. Nor are hardware developments restricted to the field of computers and associated hardware. Microform technology is moving forward and one can foresee a coupling with computer technology, as planned by Project Intrex. The dramatic development of central concern for the welfare of libraries was highlighted in congressional hearings examining recommendations looking toward the establishment of a permanent Library Commission, and thus providing libraries with an identifiable locus in our federal government. Perhaps no more dramatic development can be pointed out than the publication of the report by the Library of Congress on its study of retrospective conversion,[1] which caught the author of the article on this topic in mid-course.

These articles, even in combination, do not yield a design for a national program of library automation; nor was such design anticipated. If the selection of a topic was as judicious as was hoped for, the insights provided by these authors will be stimulating for future design efforts. Their contributions will be most useful, however, if the articles generate disagreement and controversy. It is the active engagement and concern of the library profession with these problems and issues that will ultimately pave the way toward a coherent and concerted effort that will exploit the advantages of automation on behalf of library service to which we are committed.

Reference

1. *Conversion of Retrospective Catalog Records to Machine-Readable Form; A Study of the Feasibility of a National Bibliographic Service.* Prepared by the RECON Working Task Force, Henriette D. Avram, chairman, John C. Rather, ed. Washington, D.C., Library of Congress, 1969.

Standardization Requirements of a National Program for Information Transfer

RONALD L. WIGINGTON AND JAMES L. WOOD

A NATIONAL program for information transfer must consider both the functions that various groups perform and the system that their numerous interrelated, but somewhat independent, activities constitute. Automation of many of these functions is widely regarded as the key for achieving significant advances over the present, predominantly manual, national system for organizing and disseminating scientific, technical, and other scholarly information. Automation of any activity requires that the function of each part and the multiple interaction between parts be brought into sharp focus. Usually, the first lesson learned in any automation effort is that even though the target activity may have been operating with some degree of success in the past, it is still not sufficiently well understood to enable the listing of complete and accurate requirements to guide the specification of the machine system. As a result, desired improvements are elusive and slow to develop.

The authors of this article represent two specializations in the spectrum of information transfer activities, that of the computer and communication system engineer and that of the librarian. These points of view are combined in examining the requirements for standardization in the national efforts to use automation in publication, library, abstracting and indexing, and information-retrieval activities. Standardization is necessary to both the representation of information and to the procedures being developed for handling it.

The details of standardization represent a major part of over-all system design. Underspecification risks operational failure, or at least it endangers over-all operational efficiency and economy. Conversely, overspecification can impede future improvements and extensions and can raise serious practical obstacles to cooperation.

Ronald L. Wigington and James L. Wood are with the Chemical Abstracts Service, Columbus, Ohio.

Standardization Requirements

The concepts of "network" and "service utility" have often been referred to in conceptualizing a national information transfer system.[1] Frequently it is useful to consider analogies as a source for design guidance or insight. The national telephone system was one of the earliest technologically-based networks to be subjected to scientific analysis and design, and in addition, it is one form of an information-transfer system. Standardization has played a key role in its successful development and operation.[2]

A major characteristic of the development of the national telephone network, and one that offers a valid analogy for guidance in considering the development of a national information transfer network, is that improvements are brought about by evolutionary change. Evolution in turn involves a succession of standards for each part of the system, and compatibility with the rest of the system is an essential requirement for any new part. When the investment in an existing system is very large, that system cannot be replaced or revolutionized very quickly; it must be changed gradually as the economic and human resources are available to do so. The standardization of what the system handles, and the procedures for doing it, must change with that evolution. The standardization must also be effected with great care both to gain sound current operations and to facilitate future change.

The telephone system analogy can be used to illustrate some of these points. Prior to the introduction of nationwide long-distance dialing, several different formats of telephone numbers existed, several different kinds of switching equipment were used, and correspondingly, different methods for signalling that equipment were employed in local telephone systems. Much manual intervention was necessary to establish a long-distance connection, and during periods of peak usage, performance was often poor. Before the installation of the nationwide system, a standardized method of expressing telephone numbers was established.[3] It provided a concise, unique identification of the terminal stations in the network. The information to be transferred from one point to another in the switching network was that unique identification.

Even though the initial handling facilities did not represent the information identically, ways were worked out to translate the form of that information and pass it across the internal system boundaries. And although manual intervention was used as necessary to make the total system function, such intervention was greatly reduced compared with the former entirely manual system. As new equipment has been in-

stalled over a long period of time, greater homogeneity in standards of representation and procedure has been achieved, and performance and economy have improved. Still, some heterogeneity remains, and always will, because of continuous system change. At all times, however, the unambiguous information content of the switching control signals had to be preserved and it had to be capable of being transformed into any form needed for any system function handled by man or by machine.

Standardization is far from being a new subject in information handling.[4] For example, that segment of information handling concerned with library science contains many instances of uniform practice established to enable the results of work done once to be used effectively by many and to improve uniformity of the interface to the library user.

A good example of this type of standardization is the descriptive cataloging on Library of Congress catalog cards. During 1968, the Library of Congress sold over 78 million cards to approximately 25,000 libraries, firms, and individuals.[5] By accepting in part, if not in whole, the Library of Congress' descriptive cataloging, classification, and subject headings, the library community saved millions of dollars, and at the same time provided standardized access points to its book collections. In fact, the 3 inch by 5 inch (7½ by 12½ cm.) card is itself an excellent example of library standardization.

In 1945, Wilson and Tauber in their classic work, *The University Library,* noted that "the progress of a profession is usually marked by the cumulation of an increasing number of generally accepted practices." [6] In reviewing the number of "generally accepted practices" used by the library community, it might even seem that no additional "standards" would be required for the part of the national program that involves library automation. Unfortunately, this is not the case.

First, there is a serious lack of standards formally approved for the information field by the American National Standards Institute (ANSI) and by the International Standards Organization (ISO). Furthermore, standards that have been fully approved by these organizations are applicable to only a small segment of the over-all information field. In fact, even after useful ISO or ANSI Standards have been fully formalized, there is a long delay in getting them into widespread use, usually because of economic reasons.

Second, many of the existing generally accepted practices, so often mistakenly referred to as standards, are subject to interpretation by

different individuals functioning in a variety of environments at different times and are controlled or motivated by differing conditions. An example exists in the library field in such generally accepted practices as those outlined in the *A.L.A. Cataloging Rules for Author and Title Entries*,[7] the *A.L.A. Rules for Filing Catalog Cards*,[8] and in the *Rules for Descriptive Cataloging in the Library of Congress*.[9] However, these do not lend themselves to simple, concise, rigid specifications such as one would find, for instance, in a standard laboratory technique for the preparation of a synthetic organic compound or in a mechanical dimension standard for machine screws. Records generated by different libraries for the same publication can, and usually do, differ considerably.[10] Such situations cannot be automated easily.

If one accepts the hypothesis that a national program for information transfer has as its objectives the development of a coherent system for the efficient, effective, and economic transfer of information, then the need for a standardization program which is much more extensive than anything available to date becomes obvious.

As in many other fields, automation is being introduced to handle an increased load, to provide new and better services or results, and to augment scarce human skills. Automation in information handling is being achieved through modern technological methods of reprography, data processing, and communications. Their use emphasizes the need for standardization. The power of humans to interpret information in many forms and to build their own intellectual bridges between system segments will no longer be sufficient to the task.

In fact, replacing this "non-productive" human intervention is one of the improvements hoped for through automation. Information interpretation by machine, however, requires great detail of specification and uniformity of practice. Machines can transform between unambiguous alternate forms, but they cannot "understand" and re-express information or resolve ambiguity through experience—at least the machines of today and the immediate tomorrow cannot.

A national information transfer automation program will also place added emphasis on the concept of shared resources. Few, if any, modern libraries are totally self-sufficient; neither is the library community, when viewed collectively. While libraries have managed to produce for themselves the bibliographic descriptions of their book collections, they have for the past 100 years been almost completely separated from the work of providing content analysis information for their periodical and serials collections. The need to find solutions to the standardization

problems and to perform the tasks imposed by a national program presents problems that transcend the confines of the library community and confronts instead the entire information-handling community which is comprised of the efforts and activities of authors, users, publishers, dealers, and abstracting and indexing services, as well as libraries. The standardization requirements of a national program for information transfer are affected by a wide range of diverse, yet interrelated factors.

There are many levels of standardization. At the document-handling level, microforms and reprography may be used to reduce storage volume and cost by providing selective copies to information users in place of the mass distribution of documents or the removal of source material from "lending libraries." But it will not be practical to use the technology of microforms and reprography on a national scale and reap the full potential benefits until there is a widespread compatibility at prices that are comparable with the rest of the image media, until there is equipment to produce microforms and equipment to view and copy them to full document size, and until a solution to the very complicated and serious problems involving copyright and appropriate usage charges is found. A system based on medium-to-medium recopying to bridge different forms has only limited practicability because photographic recopying cannot maintain the necessary image quality through an unlimited number of copy generations. Continued technological developments may ameliorate this limitation, but today's technology is definitely limited in this respect.

Computer-readable information-transfer media are currently receiving much attention with respect to standards.[11] The physical recording medium, and the character sets—both the graphic shape and coded representation—have been treated in considerable detail. Data formats and file organization are also being given increasing attention.[12]

The field of information transfer must not be exclusively concerned with these machine factors in the standardization required for large scale national and international systems, however. The representation of the content of the information carried through or on the transfer media must also be designed for effective machine handling. This may well be the most important area to consider for standardization. The chief problem is the avoidance of ambiguity. A secondary problem is that of acquiring the highest practical degree of uniformity necessary to make the process as economical as possible.

As for content representation, consideration must be given to:

1) forms derivable from original information via a set of rules, or conversely, from requirements placed on the originator of information to provide specific data items in specific forms (examples: author name formats, titles, and representation of dates);

2) codes maintained by an accepted authority and given specific meanings by definition (examples: American Society for Testing and Materials—CODEN, Standard Book Numbers, Library of Congress Catalog Card Number, and Chemical Abstracts Service Chemical Compound Registry Numbers); and

3) transformations allowable independent of machine handling (examples: transliteration rules from non-Roman to Roman alphabets, and handling of diacritical marks).

The President's Task Group for the Interchange of Scientific and Technical Information in Machine Language (ISTIM) recommended in their final report that:

> USASI Committee Z39 in cooperation with concerned organizations for indexing and abstracting give high priority to the development of a standard procedure for citation and that this be disseminated to the publishers of primary and secondary literature for early consideration . . . that Z39 pursue the development of standards citations to other types of materials, such as books, patents, etc. . . . and that particular efforts be made during the immediate future within the normal framework of Z39 to enhance and enlarge the participation in Z39 by the major organizations directly involved in mechanized bibliographic data handling. In this connection, it is recommended that the scope statement and name of Z39 be enlarged to encompass the related discipline of information science.[13]

ANSI (formerly USASI) Committee Z39 has thus been identified as a mechanism through which these standards of primary importance can be formulated on a national basis. The recently published SATCOM Report [14] also emphasized the broad role of Z39 in reaching agreements on bibliographic practices among libraries, documentation centers, and the abstracting and indexing services. These recommendations reflect what is actually being done by Z39, which has been supported in its efforts, since 1961, by the National Science Foundation and the Council on Library Resources.[15]

In a national program each of the participants becomes a node in a

network and the interactions between pairs of nodes can be significantly different in different parts of the network. For example, the transfer of information between a library as a node and the library client as a node may be quite different from that between a content-analysis center (such as the Library of Congress cataloging department or an abstracting and indexing service) and a library. The format of a machine-readable record is simply the framework which carries the defined elements of data or character strings within the machine record. In the context of a national program, then, the prime consideration with respect of formats should be focused on communication or interchange of information between nodes and not necessarily the formats used to process the information (store, search, or retrieve) at any specific organization. It may not be necessary, or even desirable, that the machine formats be uniform throughout the network, but it is mandatory that the content representations be uniform and that they be translatable into all formats used to carry that information.

To date, the most widely publicized format for English monograph bibliographic data distribution is that developed at the Library of Congress, the MARC II format.[16] This format is currently pending for acceptance by the ANSI as the USA Standard for a Format for Bibliographic Information Interchange on Magnetic Tape. This format is being tested and is gaining acceptance by the library community.[17] There are other formats, such as the ones being introduced in their respective fields by the Institute of Electrical and Electronics Engineers, by the American Institute of Physics, by the BioSciences Information Service of Biological Abstracts, by the Chemical Abstracts Service, and the one proposed by the Joint Agreements Group,[18] for the handling of a combination of bibliographic information and other forms of information such as chemical structures, physical data and biological properties.

Regardless of their importance to the solution of current problems, none of these present forms will, in the authors' opinion, be the backbone of the eventual national interchange program with its broad range of requirements. The success of a national program depends upon having a standard interchange format, or a campatible set of interconvertible ones, but precisely what this format, or these formats, should be is still a topic to be addressed by the participants in the national program. Moreover, some solid experience with present forms is needed now to further the development of a format suitable for broad-scale national use. The successful acceptance of any format as a

format for a national program depends on several factors. Among them are the amount of useful data made available in the format from any source, and how well it stands the test of usage.

We have strongly emphasized the necessity of being unambiguous in the representation of information content in a machine-readable record, and the necessity of being able to translate that content into any form needed for processing in any part of the information network. The approach that is being adopted very widely in the information community, for example, in MARC II, in the Chemical Abstracts Service Standard Distribution Format, and in the discussion of the Joint Agreements Group, is the definition of data elements and the labeling of those data elements as information transfer media by published codes. Data elements then are defined as units of information within the system. Customarily, a data element or unit of information contains a tag or identifier and the content. For example, in the MARC II system, the tag 245 or TIL has been assigned to the variable field entitled "Title Statement."[19] In this way, the user of MARC II formated tapes knows that a data element tag 245 or TIL contains the title of the monograph represented by the record. Similarly, in the Chemical Abstracts Service Standard File Format, data element 200 is the CODEN of the source journal.

The information community has not yet evolved to the point where these data element tags are standardized, and it is far from the point where the transfer formats themselves are identical in implementation.

In May 1967, Z39 Subcommitte 2 (Machine Input Records), Special Project on Data Elements, published a composite list of bibliographic data elements applicable to the full range of bibliographic forms.[20] In August 1967, Subcommittee 2 under the direction of chairman Henriette Avram, Information Systems Office, Library of Congress, met and concluded that:

> The most useful next step for the SC/2 would be to draft a general statement which would establish a medium of exchange between various producers and users of bibliographic information. The general principle of identification or tagging of data elements would be emphasized, but there would be no attempt at specifying all data elements to be identified in all bibliographic forms.[21]

Since that August 1967 meeting, an increasing number of machine-readable data bases have become available to the information-handling community. In science and technology alone Cohan[22] lists 289 such services, some of which are available on a subscription basis

while others are available in an experimental form only. The data-element identification work done by the Information Systems Office of the Library of Congress for the MARC II format for monographs, and more recently the working document for the MARC format for serials, have been the data-element lists furnished from the library community. The abstracting and indexing community has published, or has made available to interested parties, sets of data elements including those issued by the International Atomic Energy Agency for its International Nuclear Information Service (INIS) and by the Chemical Abstracts Service.

There is ongoing work on data-element specifications that may have a direct bearing on the national plan for automated information transfer. This is the work of the Joint UNESCO/International Council of Scientific Unions UNISIST project[23] and the National Federation of Science Abstracting and Indexing Services, Bibliographic Citations Committee (NFSAIS-BCC).[24] The ultimate goals of these two working groups is to develop an optimal set of interchange data elements that will be acceptable to the abstracting and indexing services whose staff are members of International Council of Scientific Unions-Abstracting Board (ICSU-AB) and NFSAIS respectively.

During the past two years, ANSI Z39 Committee has addressed its efforts vigorously to the data-element identification and specification problem. Its subcommittees are currently working on the following data elements: a) country names and political subdivisions (such as countries of the United Kingdom, republics of the Soviet Union, states and counties of the United States, and provinces of Canada), b) calendar dates, c) languages, d) names and addresses of libraries, e) names and addresses of book dealers, and f) names and addresses of publishers. In addition, Z39 committees are working to develop Standard Book Numbers (SBN) and Standard Serial Numbers (SSN). Standardization of each of these data elements and of many others is required for a national program.

It is apparent that a more comprehensive and coherent attack will have to be made on the data-element standardization problems than has been waged in the past. Such an effort could merge and unify the results of various national and international group efforts, such as those of NFSAIS-BCC, MARC, UNISIST, and individual and collective abstracting and indexing services, into a coherent set of community-acceptable standards. Optimal standardization will require some sacrifice of traditional practices of individual members of the

over-all information-transfer community. However, the eventual benefits of such standardization to a national program would far outweigh any restrictions on current practices or any other disadvantages it might bring about. Getting into position to achieve this standardization is a difficult task.

There are many bibliographic data fields that can be adequately expressed for interchange purposes in the forms of abbreviations or codes rather than complete data. A name of a country, for example, can be represented by a coded form on an interchange record and can be expanded to a full name or, if necessary, reduced to an abbreviation of that name on output.

The advantages of code utilization are many. Less input transcription (keyboarding) is required, and the fewer the characters, the less the chance for error. There are fewer characters to proof, fewer to correct when errors do occur, and fewer to store and transmit. Codes are also amenable to the use of check characters for error detection. Translation of the code to the assigned value has an inherent degree of flexibility. Codes become a kind of Esperanto, or universal language. For example, a serial publication might carry titles on its cover in several languages. The standard serial code for one processor would translate to the version of the language most meaningful for his clients' requirements, while for another processor the code could translate to a library catalog entry or to the title abbreviation for still another language.

Earlier in this paper the authors identified several data elements that lend themselves to coding, including place names, languages, publishers, libraries, books, and serials. Several organizations have already developed authority files for codes for their internal use and for use by their clients. There has also been some effort to standardize codes and their coded values. The recently published Federal Information Processing Standards (FIPS) include coded values for calendar dates,[25] for states of the United States,[26] and for counties within the states.[27] The International Organization for Standardization has also published a recommendation relevant to this effort.[28] The national program for library automation or information transfer will require the identification of those bibliographic data elements whose contents can be best represented by codes and then the standardization of the codes and the data they represent.

Another area that will require standardization is transliteration. The transliteration problem includes not only those languages using ideo-

graphs or non-Roman alphabets, but also those languages using alphabets that contain more than the twenty-six letters used in the English language. In 1969, the International Organization for Standardization published a revision of its 1955 Recommendation 9: *International System for the Transliteration of Cyrillic Characters.*[29] This standard is worthy of serious consideration by the program. The International Organization for Standardization Technical Committee 46 Documentation Meeting held in Stockholm during October 1969 included discussions of the transliterations of Yiddish, Chinese, Japanese, Arabic, Hebrew, and non-Slavic Cyrillic languages. Eventually these discussions must culminate in community-acceptable standards.

If direct compatibility of data-element form and content of the various systems comprising a national program is to be achieved, then the full range of the transliteration problems mentioned above must be subject to rigid standardization. However, if only convertibility is required, it can be achieved with "softer" standards. Convertibility will require that each organization producing bibliographic records make known in detail the codes, their values, and the transliteration schemes that it employs. Care must be taken to insure that these local transliteration schemes do not lose information in the process.

The following statement from the Library of Congress reflects on both the national program requirements for transliterations and the need to standardize these transliteration schemes.

> Until a definitive character set for nonroman alphabets and the techniques for input, storage, and output of such characters can be developed, all information will be entered into the serials system in romanized form. The romanization or transliteration schemes presently applied by the Library of Congress will be used when cataloging publications in nonroman alphabets or nonalphabetic languages.[30]

Transliteration problems in the library and information community are analogous to the character-set problem of the computer system field. The needs of a printer, a telegrapher, an information-retrieval system, and a data-processing system, are all different. However, each of these has an impact on the national information transfer system. True standards have been formalized only for the data-processing system that handles strings of character data (e.g., ANSI 7-bit character set).[31] However, for a national system to operate successfully, all of these interests must be brought into compatibility (transformability).

The problem of general semantics in text is so complex as to far exceed the problems of character sets and of data elements. Within a specialized field, a thesaurus is often used to provide a standardized set of relationships among terms including synonyms and hierarchical relationships.[32] It is fortunate that much can be accomplished in a national information-transfer program before tackling the problem of standardization of thesauri. This is because most of the attention at this time is focused on automated dissemination of, or access to, information, and not on the machine interpretation of it. There is much research on information retrieval concerned with machine interpretation of ordinary text. Unfortunately, this research has not yet matured to the point that it can be considered a major feature of a national information-transfer system. Furthermore, before it can be fully effective, the "data input" problem must be solved to get most relevant text into machine-readable form.

Because of the heterogeneity of the various links in a national information-transfer network and the immensity of the system relative to any change that can be made in it by any one effort, the authors have departed from the usual idealistic view of standardization as the achievement of total uniformity in the representation and handling of information. While no opportunity to achieve such uniformity should be wasted, progress will have to be achieved step-by-step as each part of the system develops.

In the analogy from the national telephone system, there existed a single management-planning and decision-making authority for most of the system involved. It could force the network design and standardization so that local, non-optimal actions could be taken to achieve over-all optimization. In contrast, there are many major centers of influence involved in the information problems:

a) national libraries and the general library community,

b) abstracting and indexing services,

c) major scientific and technical societies and their publications and conferences,

d) large mission-oriented activities of the U.S. government,

e) organizations that develop and sell communication, computer, and reprography technology on which a modern national network depends,

f) consumers of information, a large heterogeneous group of persons and organizations with diverse and sometimes conflicting needs,

g) commercial publishing and information-handling enterprises, and

h) generators of information in a multitude of forms.

Any attempt to present the above in any order of precedence, importance, or impact, is foolish; there is, therefore, no single entity, nor is there likely to be, with sufficient skills, prestige, resources, and authority to assume the role of monolithic management.

Thus, planners for the evolution of the national information network, and the developers of the standards which guide it, must take this difference into account as they strive to make the technological potential a reality. The network must grow together, forced only by common need and the recognition of all parties that cooperation and perseverance alone will lead to the desired result. The importance of the various standards organizations as the final mechanism for recognizing the transformation of "accepted practice" into formal standards thus becomes evident. And a necessary preliminary to creating these formal standards is the step-by-step experience of individuals and groups who work to solve the problems of ambiguity and incompatibility between nodes in the system.

Finally, we cannot consider the development of a national information transfer network without considering the international implications. Much of the information handled in scientific, technical, and scholarly publications in the national system originates outside of the United States. Any national system in the United States has an international impact. Furthermore, some of the organizations that will be instrumental in achieving standards, such as International Organization for Standardization and UNISIST, are international organizations. Thus, while we work toward achieving an improved national information-transfer system, we must be acutely aware of these international relationships.

As the expression of the major significant details in information-transfer system design, standards of representation and practice, agreed to and used by all parties, become the guiding mechanisms which replace unified management. As such they take on an importance in achieving progress in national, and international information transfer which is beyond the technical importance normally associated with standards. All parties, however, must be patient with the inevitably slow development and utilization of those standards.

References

1. Becker, Joseph. "Information Network Prospects in the United States," *Library Trends,* 17:306-17, Jan. 1969.

2. Nunn, W. H. "Nationwide Numbering Plan," *Bell System Technical Journal,* 31:851-59, Sept. 1952; and Adam, A. O. "Crossbar Tandem as a Long Distance Switching System," *Bell System Technical Journal,* 35:91-108, Jan. 1956.

3. Nunn, *ibid.*

4. Steel, T. B., Jr. "Standards for Computers and Information Processing." *In* F. A. Alt and M. Rubinoff, eds., *Advances in Computers.* Vol. 8. Academic Press, New York, 1967.

5. U.S. Library of Congress. *Annual Report of the Librarian of Congress for the Fiscal Year Ending June 30, 1968.* Washington, D.C., Library of Congress, 1969, p. 23.

6. Wilson, Louis R., and Tauber, Maurice F. *The University Library; Its Organization, Administration and Functions.* Chicago, University of Chicago Press, 1945, p. 82.

7. American Library Association. Division of Cataloging and Classification. *A.L.A. Cataloguing Rules for Author and Title Entries.* 2d ed., Chicago, ALA, 1949.

8. American Library Association. *A.L.A. Rules for Filing Catalog Cards.* 2d ed., Chicago, ALA, 1968.

9. U.S. Library of Congress. Descriptive Cataloging Division. *Rules for Descriptive Cataloging in the Library of Congress* (Adopted by the American Library Association). Washington, D.C., Library of Congress, 1949.

10. Truelson, Stanley D., Jr. "The Need to Standardize Descriptive Cataloging," *Medical Library Association Bulletin,* 57:21-27, Jan. 1969.

11. United States of America Standards Institute. Z39 Committee. "USA Standard for a Format for Bibliographic Information Interchange on Magnetic Tape," *Journal of Library Automation,* 2:53-65, June 1969.

12. Mooers, Calvin N. *Standards for User Procedures and Data Formats in Automated Information Systems and Networks. Part I. The Need for Standardization and the Manner in which Standardization can be Accomplished.* U.S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, July 5, 1967; Mooers, Calvin N. *Standards for User Procedures and Data Formats in Automated Information Systems and Networks. Part II. The Standardizable Elements of User Control Procedures and Unified System Model.* U.S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, Aug. 10, 1967; Mooers, Calvin N. *Standards for User Procedures and Data Formats in Automated Information Systems and Networks. Part III. A Suggested Standard Keyboard Assignment for the Elemental User Control Actions.* U.S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, Aug. 1, 1967; and Mooers, Calvin N. *Standards for User Procedures and Data Formats in Automated Information Systems and Networks. Part IV. A Standard Method for the Description of External Data Formats.* U.S. Department of Commerce, Clearinghouse for Federal Scientific and Technical Information, Aug. 28, 1967.

13. Executive Office of the President, Office of Science and Technology. *Draft Final Report of the Task Group for Interchange of Scientific and Technical Information in Machine Language (ISTIM).* Washington, D.C., 1968, pp. 9-10, 14-17.

14. National Academy of Sciences. National Academy of Engineering. *Scientific and Technical Communication.* Washington, D.C., National Academy of Sciences, 1969, p. 84.

15. Richter, Anne J. "United States of America Standards Institute Standards Committee Z-39." In *Bowker Annual of Library and Book Trade Information.* New York, R.R. Bowker, 1969, p. 249.

16. U.S. Library of Congress, Office of Information. *The MARC II Format: A Communications Format for Bibliographic Data.* Prepared by Henriette Avram, *et al.* Washington, D.C., Library of Congress, 1968.

17. Gull, Cloyd D. "Convergence toward Common Standards in Machine-Readable Cataloging," *Medical Library Association Bulletin,* 57:28-35, Jan. 1969.

18. Tompkins, Howard, E., *et al.* "Tentative Common Practices on Formats for Public Bibliographic Data Tapes," *Joint Agreements Group Report to the Fall Joint Computer Conference.* 1968.

19. U.S. Library of Congress. Information Systems Office. *MARC Manuals Used by the Library of Congress.* Vol. 2. Chicago, ALA, Information Science and Automation Division, 1969, p. 53.

20. Curran, Ann T. and Avram, Henriette D. *The Identification of Data Elements in Bibliographic Records: Final Report of the Special Project on Data Elements for the Subcommittee on Machine Input Records (SC-2) of the Sectional Committee on Library Work and Documentation (Z-39) of the United States of America Standards Institute.* New York, USASI, Sectional Committee on Library Work and Documentation, 1967.

21. Avram, Henriette D. "SC/2 Machine Input Records Quarterly Report," *News About Z39,* Nov., 1967, p. 2.

22. Cohan, Leonard, ed. *Directory of Computerized Information in Science & Technology* (International Information Network Series). New York, Science Associates/International, 1968.

23. Joint Project on the Communication of Scientific Information. *Report of the 1st Session of the ICSU-UNESCO Central Committee to Study the Feasibility of a World Science Information System, UNESCO House, Paris, Dec. 6-8, 1967* (ISCU-UNESCO/CSI/2.15). Paris, International Council of Scientific Unions, Jan. 12, 1968.

24. "Bibliographic Citation Committee," *News from Science Abstracting & Indexing Services,* Vol. 10, No. 4, Nov. 1, 1968.

25. U.S. National Bureau of Standards. "Specifications for Calendar Date," *Federal Information Processing Standards Publication,* No. 4, Nov. 1, 1968.

26. U.S. National Bureau of Standards. "Specifications for States of the United States," *Federal Information Processing Standards Publication,* No. 5, Nov. 1, 1968.

27. U.S. National Bureau of Standards. "Specifications for Counties of the States of the United States," *Federal Information Processing Standards Publication,* No. 6, Nov. 1, 1968.

28. International Organization for Standardization. *Symbols for Languages, Countries, and Authorities: ISO/R 639.* 1968.

29. International Organization for Standardization. *Revision of the ISO Recommendation R9-1955-International System for the Transliteration of Cyrillic Characters* (Draft ISO Recommendation No. 1243). Jan. 1967.

30. U.S. Library of Congress, Information Systems Office. *Serials: A MARC Format.* Washington, D.C., Library of Congress, 1969, p. 3.

31. United States of America Standards Institute, Standards Committee X3. *USA Standard Code for Information Interchange.* New York, USASI, 1968.

32. Mauperon, A. "Adaption to User's Needs," *ASLIB Proceedings,* 19:232-40, July 1967; Rolling, Loll N. "A Computer-Aided Information Service for Nuclear Science and Technology: EURATOMCID," *Journal of Documentation,* 22:93-115, June 1966; "EURATOM-Thesaurus, Indexing Terms Used within EURATOM'S Nuclear Documentation System," *EURATOM REPORT* (EUR 500e). 2d ed., Part I, 1966; Part 2, 1967.

Economics of National Automation of Libraries*

RALPH M. SHOFFNER

APPROXIMATELY 0.1 PERCENT of our Gross National Product (GNP) is devoted to libraries, yet library operations are basically the same today as they were twenty years ago. And, while library operations continue unchanged, the cost pressures are intensifying. Across all libraries, personnel costs comprise about two-thirds of total expenditures. The excellent study by Mathematica (a research organization) performed for the National Advisory Commission on Libraries [1] shows that the increase in the cost of library operations must be *more rapid* than the increase in the cost of living because there has been no improvement in productivity per man-hour to offset the cost increases. This means that there is a higher rate of inflation in libraries than in the economy as a whole; that is, the cost of living reflects the net effects of both personnel costs and productivity in the economy as a whole. In our society the cost of living has not risen as rapidly as personnel cost because there have been increases in productivity. But there has been virtually no increase in productivity in library operations; therefore, the increase in library personnel costs must cause an increase in total library costs which is greater than that of the cost of living.

Those who are concerned with obtaining the funds to support library operations are justified in their feelings of crisis. In order to provide even the same level of service, total library costs will continue to increase as personnel costs increase year by year. Adding to this the increasing demand for library services in the form of a more literate population, more students, etc., it would not be surprising if

Ralph M. Shoffner is Project Coordinator, Institute of Library Research, University of California, Berkeley.

* Partial support for research done for this article was provided under the U.S. Office of Education Grant No. OEG-1-7-071083-5068.

some areas of library service were to suffer. Even though it is impressive that total library funding has doubled within a decade, it is not clear that this can continue.

Computer-based library automation is being cited by many as signaling a fundamental change in library operations which will radically alter the productivity of the individual librarian and thus, break the direct relationship between personnel costs and the costs of library services. While accepting this point of view as theoretically sound, we should question more closely why there has been so little change in library operations in the past and whether anything has occurred that will allow the rate of change to increase.

We are in the fortunate position of being able to obtain tentative answers to these questions. In recent years, economists have been improving our understanding of the processes of technological change. It is the purpose of this article to consider library automation in the light of this understanding and to interpret its implications. The economists' work indicates that the low rate of change in library operations can be attributed to structural features of libraries and also to improper models of technological change in libraries, both of which remain largely unchanged.

Given the present organization of the library field, it appears that by 1980 an increase in productivity of perhaps 30 percent will be possible, but this increase may be in effect in only half of the field due to the size of libraries and the problems in transmitting the necessary knowledge of the new techniques. Compared with the potential for change and the need for change, this rate does not seem high enough. Therefore, at the conclusion of this paper, some problems and possible solutions are sketched which should increase the rate of change in library operations in the coming decade.

The greatest need is for a program that combines the research, development and application of automated procedures. To obtain this, some research and development groups should be identified and charged with these responsibilities for the library community as a whole. Another need which will have to be met is the increasing shortage of qualified librarians. This shortage will be heightened by the new knowledge requirements for those organizing and operating automated library systems combined with the expanding demands for library services. This shortage can be met by establishing a program of supplemental education for present librarians and for those entering the field.

Status of Libraries and Library Automation

While it is not within the scope of this article to review extensively the current status of libraries and library automation, it is nonetheless necessary to provide a brief summary of that status in order to relate it to the economic influences on technological invention and application.

Let us begin by summarizing some of the cost and size characteristics of the targets of the automation efforts—the libraries: 1) they are small organizations (in relation to industry), 2) there are a large number of them, 3) personnel constitutes the major cost (two-thirds) of library operations, 4) taken together, libraries are a noticeable portion of the national economy (approximately 0.1 percent of the GNP), and 5) libraries are growing. Consider the following data.

Industrial organizations with annual sales of less than $50 million are usually considered small. By contrast, for 1965-66, only sixty-two college and university libraries were reported to have total operating expenditures of over $1 million, and the largest of these had a total of less than $7 million. Total operating expenditures for these sixty-two libraries were $122 million. However, a total of 2,207 college and university libraries reported total operating expenditures of only $320 million. Thus, after removing these sixty-two relative giants, the remaining libraries had an average annual operating expenditure of $92,000.[2] For the same period (from 1965-66), there were 1,178 public libraries serving communities with populations of 25,000 or more. The largest of these had expenditures of $16 million. The 270 libraries serving populations over 100,000 had total expenditures of $256 million (average, $950,000), while the remainder of the 1,114 reporting (not all of the 1,178 libraries reported expenditures) had expenditures of $88 million (average, $104,000).[3]

During that same period, there were 26,500 professional staff and 45,000 non-professional staff (both measured in full-time equivalents) in all of these university, college, and public libraries. The staff costs were $405 million, or 60 percent of the total expenditure. Staff costs are about evenly divided between the professional and non-professional categories.[4]

There are still more libraries to be accounted for, especially high school, elementary school, and special libraries. Though the statistics reported are uneven, the Bowker data indicate a total of nearly 28,000 libraries of all types (including university, college and public libraries).[5] For 1968, there were library expenditures of approximately

$1.5 billion, of which $1.0 billion were personnel costs. This means that about 0.1 percent of the Gross National Product is devoted to libraries.

Finally, the scientific and technical information activities of the federal government (STINFO) amounted to about $350 million.[6] While these activities are referred to as "transfer of information," they are very like library activities but are based upon technical documents and non-book materials. If they are included, the total expenditures for library-related activities in the U.S. in 1970 will probably exceed $2 billion.

The Mathematica study[1] has provided the necessary data to project the rate of change of these library costs, given a constant level of operation. Over the fifteen-year period from 1951 to 1966, salary scales were increasing at the rate of 4 to 5 percent per year. Because of this, the cost per unit of library operation has risen at the rate of 2.5 to 3 percent per year.[7] So long as the technology of library operations remains the same, costs are likely to increase in this manner, even without changes in volume of operation. However, there are also changes in volume of operations as a result of increases in purchasing levels in response to increasing publication and increasing demand. The rate of publication is said to be doubling every fifteen years. Library patrons are placing an increasing level of demand on the libraries, and by the end of the century there will be 100 million more people in the United States. The implications of this are enormous for libraries as well as for all public services.

With respect to the book collection, the public, college and university libraries discussed above were reported to have 465 million volumes at the end of 1965-66.[8] Of this total, 34 million volumes were added during the year. This represents an increase of 30 percent over the rate of additions in 1962-63.[9] Recent data continue to reflect this increase in the rate of addition.[10]

Unfortunately, however, we are not in a position to estimate or project total demand for library services because the statistics do not report the service aspects of these libraries. For example, none of the following is collected on a comprehensive basis: the amount of floor traffic, the number of circulations of various loan periods, the number of inquiries that are made, and the number of hours that the libraries are open and servicing their patrons. One conclusion, however, can be drawn immediately: since information about current services is not available, it will be extremely difficult to make a quantitative

assessment of the service benefits to be obtained from library automation.

In the early part of this decade, the view was that computer technology would have a substantial impact upon the operations of the nation's libraries in the 1960's. While there have been many individual accomplishments, the effect on operations has, by and large, not yet occurred.

With respect to the current status of automation, Hillis Griffin has said, "Library automation is an accomplished fact."[11] Without challenging the accuracy of the comment in the context in which it was given, it should be noted that the statement allows the reader to interpret library automation as a *state* of library operation rather than a *process* of changing library operations. Indeed, library automation exists as a process and some automated procedures have been developed. Computer programs have been developed to assist with the file maintenance chores associated with acquisitions, cataloging, and circulation. Also, Brown and Jones report on the plethora of first applications of specific computer-based techniques for information storage and retrieval and information center applications.[12] However, few libraries are using any of these techniques, and there are fewer than ten libraries of the more than 3,300 public, college, and university libraries which are using them all.

Even though little application of automated clerical procedures has occurred, such techniques at least have been developed. By contrast, scant work has been done on mechanized procedures to assist professional librarians. (In referring to procedures for professionals it is assumed that library tasks divide readily into the clerical and decision categories.) No such simple distinctions exist in practice. Tasks have a spectrum of routine and decision-making aspects; thus, nonprofessionals make some decisions and professionals perform some routine activities. For convenience, however, we assume that the total volume of decision-making activity is equal to the total professional man-effort.) It is not at all clear that any present system reduces the amount of professional effort required to provide library services. Thus, the selection decisions, the cataloging decisions and the reference-retrieval decisions are made largely unaided and mostly as they have been made for years. Indeed, those organizations using automated retrieval systems have found that skilled professionals are needed to formulate the retrieval requests and to evaluate the results. Thus, the demand for professional librarians is increased. At present, one half

of a library's manpower costs are for professional personnel. As a result, research in librarianship must be addressed to assisting or replacing the professional librarian in the performance of his decision tasks. This would allow for the reallocation of the professional librarian's tasks to meet the increasing demand for library services.

Returning to the application of existing techniques, what would be the benefits and costs on a national basis of the application of current automated techniques to libraries? Because of the nature of the supporting data, the answers must be in the form of conjectures only. First of all, let us assume that the level of application we want is a computerized system that will handle most of the current housekeeping chores associated with circulation, acquisitions, cataloging, serials and management records. (Because there is no quantitative characterization of present library services, we are not at this point going to attempt to discuss the benefits of automation resulting from the possible expansion of current services or the provision of new services. Thus, the analysis is restricted to the cost benefit of automation at current service levels.)

Estimates for a library system such as the one mentioned above generally predict a 40 to 60 percent reduction in clerical costs.[13] When considering application to thousands of libraries, this savings must inevitably be cut down because it is an estimate of the best that can be done. Such an estimate assumes that when the effort requirement is reduced, the man-cost is correspondingly reduced. Yet personnel assignments might not change because of a lack of alternative work compatible with the remaining tasks that still must be performed. Two tasks, each of which is cut in half, may still require the same number of people because the tasks may require that a person be in a particular place for a specified period of time. For example, the tasks at the reference desks in two branch libraries—one medicine and the other law—would be difficult to combine satisfactorily no matter how low the loads. Similarly, the timing of processing requirements will produce interference. Thus, if telephone calls decrease, their sequence of occurrence still might be such that the person answering them could not do other work effectively.

As a result of such limitations, my guess is that on an individual basis, the libraries of average size ($100,000) or less would not notice any reduction in man-effort, while the largest libraries will have a reduction of 33 percent in the non-professional man-effort. However, processing centers represent a possible method for making the benefits

of automation real for libraries with less than a $100,000 annual expenditure. Therefore, both individual installations and processing center installations should be considered.

Examining the distribution of sizes of the public, college and university libraries, it appears that approximately 1,000 installations of a computerized system (including processing centers) would make automated techniques available to those libraries whose cumulative personnel costs total $600 million of the library community's present $1 billion personnel costs. (This is based upon an assumption that there are a large number of libraries or branches with staffs of one or two people for which installation would not provide a cost benefit.) This leads to an estimate that these libraries would realize a $100 million savings per year in their non-professional staff costs (which are half of the $600 million personnel costs). Over the entire library community, this would represent a decrease in labor costs for a constant level of operation of approximately 10 percent.

On the basis of these benefits, we can now make a rough estimate of the costs for development and installation of our computerized system on a national basis. Let us assume that the development and first installation of this system is $2 million and that each subsequent installation costs $100,000. Further, assume that one system is developed for every 100 installations and that there are 1,000 installations, including processing centers. On this basis, ten systems would be developed at a total cost of $20 million. Subsequent installations of these systems would cost $100 million, and the total cost would be $120 million.

Economic Influences on Automation

It seems reasonably clear that the accomplishments of library automation during the decade of the 1960's were much less than was expected at its beginning. It would be easy to accept the current view, which might be paraphrased as follows: At the beginning of the 1960's the problems of library automation were not understood, and they turned out to be much more difficult than anticipated. However, most of the problems now have been identified and solved. As a result, substantial automation will be accomplished in the next few years.

In effect, this beguiling approach says: "I am wiser now, so trust me." I question whether we are wise enough. After all, some automation proposals made as recently as two or three years ago have been

found to be overly optimistic. Against this background, can the estimates made earlier in this paper be trusted? How can they be checked before they are tested by history? Is there any other experience that can be used as a cross-check on how wise we have become?

There is one cross-check that can be applied: the economics of changing technology in general. Over the last decade there has been a growing interest in this field. Mansfield has produced an excellent text which brings together the diverse work that has been done in this area.[14] I have extracted liberally from this text and the interested reader is referred to it, both for extension and for qualification of the generalizations. All readers should be aware that only very small amounts of data exist to support the general assertions set out by Mansfield. Also, there is no data that indicates the degree to which these assertions apply specifically to library automation (although the effects of the assertions should be the same). Extensive data collection and analysis remains to be done to make our knowledge of technological change more precise.

In the development and use of a new technology, Mansfield defines three stages: technological change, innovation, and the diffusion of the new technique. In the first stage, the new technology must be invented or discovered. "Technological change" does not refer to a change in the use of techniques; it refers to an increase in our fund of knowledge about possible goods and services that can be produced or about possible methods and equipment that can be used for the production of goods and services. Thus, technological change, as Mansfield uses it, represents an invention or a new technique that has been developed, whether or not it is applied.

The second stage, innovation, refers to the first use of the new technique on an operational basis. Though this may be closely related to the technological change, it may also be a separate step performed at a later time in a different organization. The third and final stage, diffusion, refers to the increased use of the technique until it reaches an equilibrium with competing techniques and thus becomes an established one. Ultimately, it may be replaced by newer techniques. Later in this article, I combine Mansfield's two stages of innovation and diffusion into a single stage: application.

The rate of the first stage, technological change or invention, is increased by:

1) increases in the demand for the product(s) affected by the change;

2) decreases in the availability, or increases in the prices of the resource inputs;

3) increases in the number of people working in the field, or related fields, in a position to make the changes;

4) increases in the amount of effort devoted to making modest improvements in operations;

5) increases in the amount of resources devoted to improvements in the capital goods and other resource inputs from supplying industries;

6) increases in cumulated research and development expenditures made by the firm or industry; and

7) increases in firm size in the range of less than 1,000 employees to more than 5,000. That is, the percent of sales which is devoted to research and development increases and thus, technological change is stimulated.[15]

Libraries appear to be prime candidates for technological change insofar as the criteria given are concerned: there have been increases in demand for library services, there have been increases in the prices of the resource inputs (in the form of increasing salary and wage costs), and there have been increasing numbers of people working to make changes in the field. Why, then, has there not been more discovery of new knowledge or more invention of new techniques or equipment?

One reason is that until recently the expenditures for research and development continued at a low level in spite of the increasing demand for library service. In general, individual libraries cannot afford research; that is, the pay-off from research is usually a function of the size of the organization. Therefore, in small organizations such as libraries, the research is thought to be too risky and the pay-off too limited to be worth the portion of the library's budget that would be required. The recent expansion in federal funding of library research and development is thus extremely important to the stimulation of technological changes.

However, there are further problems in stimulating technological change, and these have to do with attitudes about the nature of the change required and the resulting allocation of these research and development funds. Technological change is characterized either as capital-embodied or disembodied change. Capital-embodied change requires new equipment (thus, capital investment) for its use, while disembodied change, in the form of changed methods and/or organi-

zation, can be applied with either existing or new equipment. Of course, many changes are mixtures of these two types of change.

The problem with library automation is that it was improperly classified as a capital-embodied change. While the computer per se is a capital-embodied change, library automation requires much more than a computer; thus, library automation is disembodied change. Setting up a computerized system requires a performance program which specifies in minute and exhaustive detail the conditions to be expected and the actions to be performed for each condition. To develop these programs and the data to be processed by them, research is required, i.e., the study and understanding of the fundamental processes of the library. Even then, the use of these programs is not necessarily evident. Education and training may be needed to make the system user aware of the workings and the possibilities of the programs.

By contrast, the bulk of the library automation expenditure has been for the application of a knowledge *that was assumed to exist.* The functions of library acquisitions, cataloging, circulation, etc., were felt to be conceptually simple, and the tasks chosen for computer applications were referred to as "clerical processes." Library automation was conceived to be a direct and single application of existing computer techniques to well-defined processes. Even now, the emphasis of the field remains upon application or demonstration and not on research. This is a reasonable strategy to promote the use of existing techniques, but it does not make a strong contribution to the development of new techniques. The changing of technology in a library (as with any other field), requires a continuing effort, one which incorporates research, development, and application.

Let us now consider the application of existing techniques (which Mansfield refers to as innovation and diffusion). The following are the determinants of the rate of application:

1) the greater the economic advantage of the innovation over older methods, the greater the rate,

2) the lower the uncertainty associated with using the innovation, the greater the rate,

3) the lower the commitment required to try out the technique, the greater the rate,

4) the greater the rate of reduction of initial uncertainty, the greater the rate,

5) as the number of firms using the technique increases, the probability of adoption by non-users increases,

6) the greater the expected profitability, the greater the rate,

7) for the same expected profit, the lower investment alternative will be preferred,

8) the lower the knowledge and coordination required, the greater the rate,

9) the less the "new behavior or social organization," the greater the rate,

10) the fewer the changes in socio-cultural values and behavior, the greater the rate,

11) the fewer the restrictive policies of relevant labor unions, the higher the rate, and

12) industry characteristics that increase the rate:

a) inclination to experiment and risk,
b) keenly competitive,
c) financially healthy,
d) durable equipment,
e) growing industry output,
f) unconcentrated industries,
g) adequate advertising.[16]

Let us apply this general information to library automation. The determinants of the rate of application show that it is reasonable to expect very few libraries to be using existing computer techniques and programs. First, the economic advantages have been unclear. There has been considerable uncertainty with respect to most aspects of the use of computers—their development and operating costs, the period to develop the operations, the stability of the computer operations developed, etc. Furthermore, because the patrons of the library do not pay directly for the services performed, increasing utilization of the library (though desirable in many ways) increases the financial problems of the library. Because libraries operate as components of some other agency which controls their budget, the use of computers does not offer any prospect of profit because the money saved is not likely to be available to the library for other uses. Indeed, the library administrator may see the threat of increasing difficulties in winning the funds needed for the library as a result of the existence of the computer operations.

Second, given the average library budget of $100,000, the relative

size of the commitment to try out these techniques has been huge. This commitment is large compared with the budgets of even the largest libraries. Third, where the attempt has been made, the usual result (after some effort and delay) has been that the existing computer programs embodying these techniques could not be used or were unavailable. Where computer programs already exist, their operation must be understood and perhaps modified, arrangements must be made for their operation on a specific computer, and the personnel must be trained to provide the needed data and to use the results. To date this has led to a slight reduction of the initial uncertainty about the advantages of computer operations.

Whether or not the library administrator is inclined toward experiment and risk, libraries are in a position in which the penalties for failure for both the library and the administrator are far greater than have been the possible rewards from the successful application of computers to library operations. Here again, federal funding has been a most important factor in stimulating application because it provides the development capital and thereby reduces the risk to the library. However, this risk capital in general has been provided only for the initial development and first use of a system. Thus, unless the uncertainty with respect to the costs and benefits of the methods is reduced, it can be expected that there will be a low rate of subsequent application in other libraries.

Problems and Prospects

In looking forward to the decade of the 1970's, there seems little question that automation will indeed have an impact on the operation of the nation's libraries. It is also likely that the rate of change of these library operations will be greater than it has been over the last decade. The reason is that there is an increasing understanding of the problems of library automation, and some of the useful techniques that have been developed will be generally applied. Even so, most of the economic factors which contribute to a slow rate of change are still very much present as a result of the structure of the library community. Before we can predict their precise influence, the quantitative data on current performance must be collected, and quantitative analysis must be made to identify effects of each of the factors. This will require both time and effort, and, for the present, the question is whether our current approaches to library automation can be improved. My view is that there are several solvable problems which

would improve the rate of library automation, but they require regional or national coordination. I will discuss the problems and the possible solutions, but not the mechanisms for achieving the needed coordination.

When library automation moves into an area that is clearly new, such as the mechanization of professional processes, support for research and education must be corollary parts of the program, with corresponding increases in the funding. For example, in *Technology and Libraries*[17] the costs of a library improvement program are projected, and the allocation to research, education, and training is half of the amount allocated to hardware (i.e., computer equipment) and software (i.e., computer programs) specification and development. This sort of estimation is most welcome since it signals recognition of problems peculiar to librarianship which will have to be addressed as a part of the continuing program of library automation.

Because even the largest libraries are still in the category of small industrial organizations, the funding for research and development will have to continue to be external, presumably from the federal government. Therefore, some number of organizations should be selected to receive the funding for library research and development efforts on a continuing basis for periods of not less than five years. In return for the external funding, arrangements must be made to make the results generally applicable and available to the libraries. This is discussed later.

I suggest further that the funding go to a number of organizations, rather than just one, because parallel research and development efforts have been shown to be effective where there is uncertainty about the best technical alternative to be developed.[18] In library automation, there are many situations in which the "best" alternative will be unclear. If there were only one research and development group, chances are that we would not be given any alternatives to the work of this group. However, if there are several organizations doing parallel work, we can pick and choose from a variety of proffered alternatives.

With respect to the kind of research to be performed, the next stage of technological change (i.e., the one following the mechanizing of clerical operations) in libraries must assist or replace the present activities of the professional librarian. Only in this way will we be able to expand library services to meet the sharply increasing demands. Therefore, this funding must support research on the fundamental problems of librarianship, such as the intellectual organization

of information, as well as the more technically oriented problems directly related to computer operations. Indeed, the richness of these problems being worked upon as an integrated set gives the promise of significant technological change for libraries during the coming decade.

While the establishment of a set of stable research and development groups should provide for continuing discovery and invention, it does not address the problems in the application of these techniques to operating libraries. The most important problem here is that of uncertainty with respect to the *merits* of the application and to the *methods* for introducing it. In approaching this problem, the research groups should investigate the economic aspects of their work on a continuing basis. To support this work however, cost accounting data must be available from the libraries. The availability of this data would allow study of the economics of current techniques in relation to new techniques. In contrast to the many speculations in this paper, this data would provide solid information on which to base the research, development, and application of new techniques.

The research and development groups should be organized so that they provide continuing support to the library community in the application of the techniques developed. In addition to providing programs and written materials, they should train library personnel and consult with them on problems of implementing and operating library systems. Such contact is vital to insure the application of developed techniques in many libraries. Because most libraries are quite small, they can automate their operations only if they are given this kind of support. For those libraries that elect to use processing centers (instead of buying and installing their own equipment), the research and development groups should train and consult with the processing center personnel and provide support for training the library personnel.

One problem we should anticipate in trying to accelerate the rate of technological change is the potential negative reaction from library personnel. Even though total demand for library services continues to increase, there is a possibility of dislocations and/or instability of employment. As indicated before, the next series of technological changes must be addressed to assisting or replacing the activities of the professional librarian. As such it has the potential of partially obsoleting the education and experience of present librarians, while it creates demands for different education and experience.

It should be possible to establish security of employment within

the profession as a whole because of the increasing total demand for library services and because the librarian's education and experience will be only partially obsolete with national automation of libraries. A promising approach would be to establish a fellowship program to support any and all librarians for re-education in computer-based library systems. In addition to classroom work, such a program should also have an internship devoted to working with automated procedures related to the individual's particular specialties. This program would consist of many different aspects and courses of study, and it is not expected that everyone would take the same set of studies.

What are the possible dimensions of such a program? We might assume that one year is the longest time that people would take in the program, while the shortest would be two months. Now, it is not likely that all librarians would elect such a program; therefore, let us assume that over the next decade 30,000 people (out of an estimated 50,000, including those entering the field during the decade) undertook such a program. Further, let us assume that one-half of them took a three-month program, one-quarter a six-month program, and the remaining one-quarter took a full-year program. This would total 15,000 man-years of student time. Thus if the program were set up to accommodate 2,000 man-years of student time per year, this program could be accomplished in seven and a half years. If the cost were $40,000 per man-year for the support of all aspects of the program, the annual costs would be $80 million.

While this is a significant amount of money, it represents less than 10 percent of current library manpower costs or less than the amount that these costs can be expected to increase over the next two years. Since total annual costs for library services can be expected to exceed $2 billion, it would be reasonable to invest this much in library personnel training in order to get on with the tasks of technological change.

References

1. U.S. National Advisory Commission on Libraries. *On the Economics of Library Operation. Final Report Submitted to National Advisory Commission on Libraries* (ED 022 523). Office of Education, Bureau of Research, Washington, D.C., 1967.

2. American Library Association. Library Administration Division. *Library Statistics of Colleges and Universities, 1965-66.* Chicago, ALA, 1967, pp. 88-162.

3. U.S. Department of Health, Education and Welfare. Office of Education. *Statistics of Public Libraries Serving Communities with at Least 25,000 Inhabitants.* Washington, D.C., 1965, pp. 1,3.

4. *Ibid.*, p. 4; and American Library Association . . ., *op. cit.*, pp. 6-9.

5. *Bowker Annual of Library and Book Trade Information.* New York, R.R. Bowker Co., 1969, p. 7.

6. U.S. National Science Foundation. *Federal Funds for Research, Development, and Other Scientific Activities; Fiscal Years 1967, 1968, and 1969* (NSF 68-27). Vol. 17, 1969, p. 65.

7. U.S. National Advisory Commission on Libraries. *On the Economics . . ., op. cit.*, pp. 13, 14, 16.

8. American Library Association, *op. cit.*, pp. 12-85; and U.S. Department of Health, Education and Welfare, *op. cit.*, p. 3.

9. U.S. Department of Health, Education and Welfare, *op. cit.*, p. 3; and American Library Association, *op. cit.*, pp. 6, 7.

10. U.S. Department of Health, Education and Welfare. Office of Education. *Library Statistics of Colleges and Universities; Data for Individual Institutions.* Washington, D.C., 1967.

11. Griffin, Hillis L. "Automation of Technical Processes in Libraries." *In* Carlos Cuardo, ed., *Annual Review of Information Science and Technology.* Vol. 3. Chicago, Encyclopaedia Britannica, 1968, p. 259.

12. Brown, Patricia L. and Jones, Shirli O. "Document Retrieval and Dissemination in Libraries and Information Centers." *In* Carlos Cuardo, ed., *ibid.*, pp. 263-88.

13. Becker, Joseph and Hayes, Robert M. "Handbook of Data Processing for Libraries." (Review copy. Publication by John Wiley and Sons expected early 1970.)

14. Mansfield, Edwin. *The Economics of Technological Change.* New York, W.W. Norton, 1968.

15. *Ibid.*, pp. 10-33.

16. *Ibid.*, pp. 119-23.

17. System Development Corporation. Information Systems Technology Staff. *Technology and Libraries.* Santa Monica, Calif., System Development Corporation, 1968.

18. Mansfield, *op. cit.*, p. 71.

Libraries, Manpower and Automation: Shaping the Future of Libraries*

MARY LEE BUNDY

BY 1980 THE TREND toward national library programs was sufficiently advanced so as to decide the role of libraries in the total information complex. What had for some time been a confusing and cluttered scene straightened itself out, although complicated by the period of intense disorder and disruption in the American society generally. As we shall discuss, this social disorganization proved first to be a deterrent and then a stimulus to information development.

The two elements—centralization and automation—did have a most decided impact on the library scene. Developments which had previously been gradual were vastly speeded up in the early seventies. Bibliographic control of the printed literature was accomplished with national and regional programs together providing not only ready access to cataloging information, but also taking on the major share of the tasks of acquisitions and book preparation.

The most resisted aspect of centralization on the part of the university research library, as well as smaller academic libraries, was in the area of cooperative acquisitions. Cooperative programs improved knowledge of the whereabouts of items, and these "networks" undertook to use the newer communication technology to speed the transmission of materials. Yet libraries still were loathe to give up the ideal

Mary Lee Bundy is Professor, School of Library and Information Services, University of Maryland, College Park.

* The writer is indebted to Paul Wasserman for his contributions to the analysis portion of this article and to Francis Lévy for insight into future information work roles.

of local self-sufficiency and the notion that collection size and service capability were synonymous. Cooperative efforts were hampered by accreditation standards for libraries which stressed collection strength. Only quite belatedly were these standards revamped to substitute service capability measures for collection evaluation criteria.

Efficiency studies showing the greater economy of obtaining little-used materials from central sources over the cost of acquiring and maintaining materials did receive acceptance with the medium and small academic library. But it was not until the active intervention of several college presidents that inter-library acquisitions programs were forced on the major university libraries. There was then quite rapid acceptance of the fact that no library could hope to build a definitive collection on any subject, except for a national client group.

This era then saw the working out of cooperative acquisitions programs of more than a token character. The evolving pattern varied with the state and the region and followed precursor efforts, such as joint storage centers and assignment of subject responsibility to large libraries. The smaller research library also won considerable recognition and the "right to serve" many constituencies in smaller sub-disciplines. The competition among libraries was a lively one with user groups joining the fray.

The resistance to microforms also gave way for several reasons—a major one being the easy ability to make full-size copies. Part of the acceptance of this alternative was undoubtedly provided by the library lootings of the early 1970's. In all types of libraries, whether regional or local, routine, repetitive tasks were automated. What had been an initial resistance to automation became a landslide. The next layer of library tasks, what might be labeled semi-professional, was transferred to the newly developed technician class. These developments—and the increased use of media other than print—have resulted in local libraries operating with a fraction of former staff, collection and space. While reducing their size tremendously, libraries have vastly improved their ability to supply known items on request.

The library profession thus came out of the 1970's with two of its traditional functions intact—bibliographical control of book and serial titles and the "supplying" function. (This was not uniformly true. Some information centers have undertaken to supply materials for their client group as a natural corollary to their information service, but most rely on library subject centers for this purpose.) In one sense, what transpired might be described as the magnificent decline of American academic librarianship.

Information center developments of this period followed paths set by early pioneers who assumed the task of information processing. At the time of this writing, they are growing rather than decreasing in number. But the most difficult task, that of coordinating effort, has largely been achieved, and previously unserved groups now have information access.

The general pattern for information centers is national centers linked to each other and to local information centers. Some of these are maintained by local groups such as universities; some are regional outlets of particular national centers. An integral part of this system is the community "interpreter" who functions both out of a local information center and out in the community served. Commercial firms still play a role in information access activity by providing specialized service to one or another group, particularly mission-oriented groups. There are a few libraries playing information center roles, undertaking S.D.I. and other user services, but by and large, libraries coexist with media centers and information centers in the city, schools, and universities and in government and industry.

One reason for this separation, despite administrative efforts to consolidate units, was the exodus of more activist-oriented librarians from conventional libraries into these newer outlets. Gradually efforts to mix passive and active elements were abandoned; it was conceded they simply did not mix.

Library schools underwent interesting developments before their exact role was determined. For a while they became the battleground between the traditional and progressive elements in the library field. There was a temporarily active group who sought to give the library schools a permanent and major role in preparing information workers for local and national centers. They sought to effect a variety of changes, such as shifting accreditation of library schools so as to involve newer professional societies and schools. These and other efforts at coordination were largely unsuccessful, partly because they were launched too late. Again, the elements, despite librarians who were active in information work, simply did not mix. The library field was largely resistant, and the information science groups indifferent.

What occurred instead of a merger was the establishment of schools of information science at most of the major universities, using as a model the early schools at Georgia Tech and Lehigh. Two library schools did transform themselves into schools of information science. The power struggle was then between the information storage and

retrieval interests and computer science departments. A number of events explain the eventual establishment of information science as a separate discipline. There were a number of "break-outs" from computer science schools, faculty who either joined or started information science schools. These schools had the greatest success (the library schools were never seriously in the competition) in producing research scholars. As a consequence, their identity and their preemption of the field was made a reality.

There was steadily increasing pressure from the national and local information centers for professional information workers, and it was not long before these schools undertook not only the research but also the professional commitment. Beginning as they had with a research orientation, it proved easier for these schools to maintain their science base than it had been for the library schools seeking to transform practical programs into theoretical ones.

We cannot account for the large enrollments which these schools have enjoyed, except that perhaps an emerging profession was of particular appeal to young people coming out of undergraduate programs during this period. Black students, especially, identified their social commitments with it. This early popularity put these schools in a good competitive position with the other professions. From the outset this has been a male-dominated profession, although from the beginning there have been and are outstanding women associated with it.

What happened to the library schools? The combination of technicians and automation took away a large share of the market for their product. The addition of subject specialists and systems analysts to library staffs closed other prospects. A large proportion of their students had been from the school field. With the development of educational communication programs, the schools of education undertook and won renewed recognition as being the place to prepare information media workers for schools. (Interestingly, these schools are now in a period of competition with the information science schools.) But the library networks still have managerial roles to be played. It was not surprising, therefore, that several library schools elected to become a department in schools of business and public administration. Other business schools developed similar minors in their programs.

In the late 1970's the Library of Congress became alarmed at the fact that there were now only a handful of library schools still open and capable of education in the bibliographical expertise required to

maintain their bibliographical system. In 1980, they established a national training program to educate for bibliographical roles in the Library of Congress and in other libraries. (The modernization of the Library of Congress is an exciting story in itself.) Dwindling enrollments, reductions in positions, retirements, and finally the closing of library schools marked the end of the efforts of librarianship to maintain its educational programs in the university and at the graduate level.

To illustrate the dynamics of the field during this era, we mention several other events and movements which, while largely unsuccessful, did serve to bring the issues into the open and influence the career choices of the people involved.

It looked for a while as if a young group of librarians—Librarians for Social Action—might shift the library picture. They sought to work in two spheres—the city ghettos and the universities. They were very much caught up in the turmoil of the 1970's, and, while they did not influence the traditional library to shift its orientation, they were moving forces in support of powerless elements and brought to the fore their demands for "information rights."

One reason this group of people failed is that they did not add to their numbers appreciably. The library schools did not make the major shifts in admissions and recruitment and programs which would have been necessary to attract the activist type. Indeed, the records of the time document the disillusionment of librarians who recruited from their communities and sent people to library schools. Their bitterness was greater because it was these very schools whose faculty had encouraged them as students.

We would suggest that the central reason why this group failed was their stubborn insistence, in the face of all evidence to the contrary, that libraries could be reconstituted so as to be socially responsive. For instance, they wasted much effort which might have been better spent in forming new associations in trying to influence change in the American Library Association. Similarly, they sought to work from within local libraries instead of transferring their operation completely outside the library milieu. These efforts did, however, make for lively episodes at national library meetings.

Social action did, however, become a major component of the information profession. The schools of information science did take on a social conscience and commitment. This came about for several reasons, but primarily because of the influx of behavioral scientists

and professionals from other disciplines who joined the faculty of these schools. They in turn attracted a different type of student and generated the new information "breed." The combination of technical competence and social commitment has been the basis for most of the successes which this field has enjoyed.

The part played by the unions during this period should also be noted. Unionization did improve working conditions in libraries and so was a force of some consequence for improvement. But it also acted to protect the senior staff. While tasks were transferred to the computer, obligations to the senior staff were maintained. A few unions did undertake to fight on professional issues. There were even remarkable instances of the union taking over the management function of trying to secure financial resources for the public library. Union activity was not a major force in the end because it occurred at a time when many public libraries were being closed as the result of schools taking over children's services. In one sense, decentralization and the building of "new communities" proved an impetus to public libraries. Almost uniformly, these communities established a popular library, staffed by local personnel. While it serves no major community, information, or educational role, it nevertheless survives on the American social scene.

We believe the information profession has become all that its most ardent supporters could have hoped for. American society has come to place a high value on information access. Tremendous strides have been made in solving technical problems, while research potentials still abound. But its success does not lie in the technical realm alone. We also believe that most of the political and social problems associated with information access have been resolved. It has proved possible to have a viable information-producing and -processing industry. The consumer's interests are protected in several ways, including provision of government inspection in those areas where an "information monopoly" might be said to exist. The legislation which brought this about, and incidentially solved the copyright problem of earlier years, is frequently cited as a model of the new relationship between government, industry, and the consumer.

The information profession can be proud of its record in the period of disorder, particularly as it had to solidify itself and at the same time make contributions toward resolving the social crises of the times.

We might cite successes in consumer terms, although these will be well known to the reader. In 1970, faculty and students at universities

had virtually no information service. Public libraries were providing service only to the reading elements of the white middle class and then mainly for recreational reading of a superficial innocuous character. It proved perfectly possible with automation, management expertise, and national information support to provide the entire range of information services to students as well as faculty. Sophisticated identification of city elements in information terms led to local and national services which did reach Americans in their occupational, cultural, social, and political lives. Information access is not only a possibility; information utilization is a reality. Children growing up today find it hard to believe the earlier pre-history of information deprivation.

This is, of course, merely a story. It did not happen, and will not happen, at least not like this. Why did we write it, particularly for a *Library Trends* issue concerned with automation and networks and in an article designed to analyze manpower implications? Because we did not accept the "givens" in the situation. We do not conceive the manpower question as merely how to be supportive of current trends or even what professional leadership may deem to be desirable professional goals. Technology and centralization are merely means; their utilization needs to proceed from socially deprived ends. We must ask the question, automation for what?

Of what value is a futuristic projection? We are sure there are many in and out of librarianship who would not quarrel with, indeed would find acceptable, the future role we posited for libraries. What they would not like, however, is the consequent loss in professional power and prestige which would accompany this relegation to the custodial function in information activity. Further, we believe many in this field are not yet prepared to accept the fact that real information needs exist and are not being met by libraries and will eventually be met in one way or another. We also believe at this point in time no corner of the profession should be sacrosanct. This is then an assault on traditional views of libraries.

But, we hope, despite the sketchy nature of our projection, it can be of positive help to those committed to insuring libraries an important part on the information scene. If it helps to identify key forces and decision points which will decide the direction of the field, it can assist in efforts to insure a future quite different. When we broaden

our perspectives beyond the confines of the past or the "here and now," we can see needs to be met, problems to be overcome, and potentials to be capitalized on. An important point to be made is that people are not at the mercy of irreversible technological and social forces. We have not described an unalterable sequence of events. Those in and out of librarianship, whether by action or inaction, are going to decide the future course of libraries. It is to be remembered particularly that if something occurs even once, it is within the realm of the possible.

Our information profession was an idealized model designed to introduce elements which we believe should characterize its development. We do not really have any basis for suggesting that the technologists will concern themselves with social goals in an active, positive way. Indeed, this has not been the history of the engineer. As Merton has expressed it:

> Deriving in part from the specialization of functions, engineers . . . come to be indoctrinated with an ethical sense of limited responsibilities . . . .
>
> So, in many quarters, it has been held absurd that the engineer should be thought accountable for the social and psychological effects of technology, since it is prefectly clear that these do not come within his special province. After all, it is the engineer's "job" —note how effectively this defines the limits of one's role and, thereby, one's social responsibility—to improve the processes of production, and it is "not his concern" to consider their ramified social effects.[1]

We do not note in the information science literature or the curricula of the new schools much evidence to the contrary.

The social events we hinted at are pure speculation, included only to make the point that library and information activity to be relevant must be viewed in a social environment which at this time is one of turmoil and violence, and we believe that efforts to respond positively to social unrest and strife may help libraries find their purpose and identity.

Our characterization of library developments does grow out of analysis of the current library scene. And since we are asking the reader, if he accepts the possibility of this future for libraries as even probable, to become an advocate for change, let us look at just how plausible our projection is.

The Current Scene

Automation. We believe the exploitation of automation and centralization to improve existing library practice is well underway. What first was a resistance to automation has apparently become a landslide for it. We find the profession in 1970 very much absorbed in employing the computer to perform traditional tasks in libraries. Undoubtedly many programs are more talked about than a reality, and much effort is still at the most primitive trial-and-error level. Yet there is clear evidence of acceptance of the computer in libraries as illustrated by the establishment of the Information Science and Automation Division in the American Library Association and the inauguration of the *Journal of Library Automation.*[2] Its pages document how libraries have come to find ways to utilize high-speed machinery to perform library tasks. The computer and its associated contribution, systems analysis, are being employed to do it more quickly, more cheaply, or more comprehensively. The modernizing of the Library of Congress, set in motion by the King report,[3] continues with the work of the Information Systems Office. Its principal product so far is the MARC project whose contribution and utility is clear when we view the time-consuming and costly work of preparing copy locally.

Library Systems. Though it is hard to separate the claims and the publicity from the reality, systems development seems also to be taking on a new lease. In the past, leadership in regional library development has come from the public library interests. Now we find academic libraries moving more strongly in this direction with the accent on cooperation among all types of libraries. Many of these programs have not gotten beyond the generation of a union list of serials, and most academic libraries are probably still not prepared to commit themselves to such programs to the extent of abandoning their own collection-building aspirations. While many efforts must be characterized as grudging and limited, imaginative programs do exist, such as the Ohio College Library Center. Public libraries are probably to be most complimented for success in getting behind-the-scenes tasks done cooperatively. We are not convinced, however, that cooperation has made any basic change in most local public library outlets, many preferring to let cooperative arrangements permit them to continue to serve recreational reading functions with the more specialized questions and demands channeled elsewhere. Yet we know several states where the responsiveness of public library systems far outpaces college libraries.

At the national level, the National Library of Medicine, with its MEDLARS project, stands out as one of the few libraries engaged in the entire range of information functions associated with serving a national audience. As we shall see, the more ambitious and more enterprising national systems are arising from other quarters, because, we believe, library leadership still sees itself tied to traditional goals and approaches.

Internal Organizational Adaptation. Internally we find libraries are not only engaged in automation efforts, but that larger libraries are also employing other specialists—systems analysts and subject bibliographers in the academic library. Public librarians are very much absorbed with issues regarding library technicians.

We suspect, however, that these developments are being inserted into libraries so as not to upset the organizational status quo too seriously. Size is still seen as a desirable goal in itself. Academic libraries still absorb the major share of their staff and other resources in acquisitions, maintenance, and control functions. It is because the absorption is here that client relationships so frequently reach a low ebb. It is not just the student but also the faculty who react against the essential client indifference of the academic library. The large libraries in particular seem bound by size and other commitments leaving them unable to adjust and adapt, while the most imaginative innovations are coming typically from smaller and newer institutions.

Libraries in a Cultural Context. In the two most turbulent sectors of society today—the city and the university—we find library service to be most in jeopardy. There is aroused concern among those in public librarianship because of its essential irrelevance in central cities. While the suburban library is more prosperous, there is evidence that it is serving little more than the ephemeral reading interests of only a small proportion of its total population. For the man in the street, the library has become a symbol without meaning; for poverty elements, it is unknown and unused.

Unfortunately, we do not hear the same soul-searching on the part of academic librarians. The "self help" notion with regard to student service still permeates professional thinking, and the library is now becoming the target of attack by students who may perceive it as an element of the older order, not attuned to their needs and aspirations. The symbol of this non-response may be the impersonal file of cards which in several instances has been the object of violence.

We know that in school, industrial, and governmental settings, con-

ventional libraries frequently coexist with newer information agencies. And though we cannot explain the dynamics of the formation of separate agencies, we believe much of the blame can be laid at the door of the conventional library for failing to respond aggressively to new potentials and new challenges.

Power and Politics. Another force is being inserted into the local public library scene—that of the unions. Staff associations, like relics of a former age, are apparently giving way to the more potent force of the union with its stronger bargaining instruments. The library profession has been delaying judgment on these efforts, seeing all the undesirable aspects of unionization, particularly in the light of recent union activity in New York City. Some library unions do appear to have pursued professional goals, and we cannot argue with their efforts to correct the bureaucratic ills. The danger seems to be that in a time when libraries should undergo drastic overhaul, the unions, with their emphasis on security, will act ultimately to rigidify them still further.

On the local political scene, we find public libraries in some cities the recent target of extensive cutbacks. While the public response may seem reassuring, the appeals still seem to be coming from parents and from nostalgia. Nowhere is support heard for the public library at these times in terms of its role in information dissemination. The possibility that schools may take over the children's trade is looming before the public library in at least two states.

Nationally too, at the time of this writing, library interests are faced with setbacks and curtailments. The report of the National Advisory Commission on Libraries is likely, like most politically inspired documents, to disappoint those who thought it might serve as a force for change. In the American Library Association there is the first outspoken dissent in many years. Those concerned with the social responsibilities of libraries have won at least an organizational foothold in ALA.

Education. There is movement in library education in the direction of adding newer information specialties to library school faculties, notably at Berkeley and Chicago, but also at other schools. The impact to date has not been extensive, and probably most library schools are not even at this stage of change. Maryland is still pretty much alone with its strong behavioral orientation.

There are voices in and out of library education asking for more

active, more aggressive students to be recruited in the field, but we do not see this concern translated into concentrated, sophisticated recruitment programs designed to achieve this goal. Student activism is beginning to be a factor in library education, but recent ripples, such as the student-organized Congress for Change, cannot be taken to mean that the library school product yet varies appreciably from that of the past. And activism will only have a point when combined with a meaningful professional preparation. If there is genuine concern over these and other issues confronting library education still oriented toward traditional courses and traditional modes of instruction, it has not reached the pages of the library journals.

The information educational scene now has not only the early pioneers, Lehigh and Georgia Tech, but also such interesting newcomers as the Stanford Communication Program. Meanwhile, information retrieval is becoming a component of computer science programs. Several research efforts will bear watching—Parker at Stanford, Salton at Cornell, MIT's Project Intrex, and what Hillman at Lehigh and Maron at Berkeley are generating. The interesting curriculum work goes on through the American Society for Information Science, while the American Library Association still keeps its exclusive accreditation rights. There is still no clarity as to where the information preparation of the future will be lodged, but the educational split is now a reality. And nowhere is there yet a coalescence of technical concerns with social issues.

The General Information Scene

There is movement in the information world coming from many quarters. As we view these developments, we look for evidence of library participation. COSATI, the Committee on Scientific and Technical Information, continues to address itself to pertinent information questions. Library interests here seem more frequently overlooked than not. While the impact of EDUCOM which was established by the Interuniversity Communications Council, is not yet clear, this instrument, with its concern about cooperative efforts and its politically sophisticated organizational structure which allows for the influence of academic power through councils and conventions, should be a factor strongly considered when calculating the educational information equation. Although it concerns itself with what are clearly library issues, there may be some information fallout through library participation. The Neighborhood Information Center movement is

apparently underway without library involvement, while some Model Cities programs are apparently involving public libraries. Other public libraries have, however, turned these planners away. That public libraries are not more concerned at these "market" losses means simply that they do not see themselves offering service of this nature.

The scientific societies are the scene of much information work because of their long-standing interests and because of the influence the National Science Foundation has been exerting in this direction. Publishing firms are taking an increased interest in information; new companies and new complexes are almost a daily phenomenon on the information scene. The information industry is beginning to take shape to the extent of having a professional association with a Washington office speaking for its interests. Their contention is that no longer should the non-profit organizations—government, universities, and societies—maintain a monopoly over research, development, and the production of information products.

The large-scale activity of consulting and quasi-research organizations who attempt to devise systems and plans for military and technical government organizations is another phenomenon productive of new forms and new patterns of information control. Many of these same interests are involving themselves in analyses of libraries.

These and other developments are important in estimating the future of information activities and the capacities in which various types of agencies will be serving. What can be said generally regarding the extent of library participation? It must be seen to be present but weak, because traditional perceptions of libraries are being projected and because library interests are not pressing seriously for important portions of the information pie. Various interests looking for information support dismiss the library and go on energetically to calculate alternative ways to get service.

There are those in and close to librarianship charging library interests with the potential loss of information functions. In an article encouraging special librarians to seek expanded functions for their agencies, Herbert S. White, recent president of the Special Libraries Association, ends by saying:

> We can fight to retain what was ours by default at a time when it was too mundane to interest others, and has now become a challenge of tremendous scope which has attracted many outsiders —some earnest and qualified, some quacks and charlatans. We can fight to demonstrate to others what we so clearly know, that the

management of information services is properly ours by training, experience and attitude. Or we can nestle securely in our fortress, ordering material only on demand, indicating its location in the system once it arrives—through an intricate cabala of symbols—and keeping accurate records of who borrowed what. This is a job even our newly arrived competitors in the information business are willing to concede to us. After all, who wants to spend his life running a stock room? [4]

Ferdinand Leimkuhler, writing in *Wilson Library Bulletin*, has placed the issue squarely before the academic library:

In the meantime, the research libraries will of necessity add more materials in microform to their collections and will make progress in the computerization of their routine operations. But the major breakthroughs may come outside the university libraries, among the government-sponsored information centers and under the aegis of the scientific and technical societies. If this happens, researchers may "plug in" to the new systems from their own offices, and the library will be eclipsed as the repository of information.[5]

Carlos Cuadra, principal investigator of the report on *Technology and Libraries* commissioned by the National Advisory Commission on Libraries, has said:

It is in no way necessary or inevitable that libraries shift the balance of their holdings and services to include microforms, digital information, videotape, holograms, and other trappings of advanced technology. It is not necessary that libraries shift their concept of operations from circulation toward outright distribution. . . . It is not necessary that libraries become elements of networks for the rapid identification and provision of material to users, regardless of geographical location.

However, these functions are going to take place; and if the library does not bring them about, some other type of agency will. That agency will then occupy the central role in the information business—the role that was once occupied by the library.[6]

If the library profession cannot be encouraged to move because of this potential market loss, it might move in response to the needs of its clients. Robert S. Taylor finds that the library's traditional passive role and its book orientation stands in the way.

As a result of the concern with books as physical objects, libraries have become static institutions concerned with the techniques of

materials handling. Whatever the reasons for this may be, and there are many legitimate ones, the library has nevertheless changed from a humanistic institution to a supply depot concerned with inventory and control. The processing of objects has blinded the library to potentially more dynamic roles as a major channel of communication, and a major processor of knowledge, both factual and fanciful, in all media. This implies not so much a change of function as a change in attitude.[7]

Having no strong client commitment, no clear-cut notion of its community responsibility, the library profession is silent when it should be speaking out loudly, passive when aggressiveness is called for, and indifferent when aroused concern is required. Great imbalances in information access continue to exist because groups lack power to change the status quo. Where libraries as public institutions might speak out for neglected constituencies, they remain value-neutral and passive. And so they must be seen to be the servants of power rather than the defenders of public rights. Intellectual freedom does continue to be defended by a few, but this is basically a defensive position. Positive identification with the information rights of people is not yet within the realm of this profession's consciousness.

The situation described above pertains to the local library situation and nationally. The legitimate spokesmen for the profession are silent on key information issues. The rallying point and the protective mechanisms for a professional group do not exist at this point in time. The image—and the impact—of a socially responsible profession with an organized point of view on key issues are absent.

Hope and Promise

There are then failures to respond to the potential of the computer, to potential market losses, and to the needs of clients. The gap beteen library activity and information activity continues to widen. What promise is there? Despite the general library situation, the current scene also reveals potentials for change. Dissent and discussion, as we have noted, while still modulated, are louder than we have heard before in this profession. The discouragement of librarians we meet and talk with can be viewed as a positive element, for out of discouragement can come support for programs of change. We have made mention of a concern on the part of librarians that libraries take on social point and purpose, a notable example being the formation of the Social Responsibilities of Libraries Round Table. There are those

within the union movement seeking to give professional purpose to this alliance with outside interests.

Promising departures in practice do exist—for example in academia at Stanford, Hampshire, and Federal City College, and undoubtedly there are others not as well-known. Public librarianship is beginning to produce a corps of workers dedicated to helping the disadvantaged. It is probably only a matter of time before they articulate the bureaucratic and legislative support needed to support poverty efforts. Library education here and there is percolating new ideas, and new types of teachers and specialists are being added to their faculties. These faculties may well become the rallying point for the activist students in their programs. We might view the present automation of libraries as a preliminary "housecleaning" preparatory to taking on more active functions.

All these offer hope that the profession has the capability for change today. And there is still time, for information activity is not all that merged. No serious claims are being made by any one group to control information practice. Information opportunities abound everywhere but particularly in the city and in academic settings where the conventional library is yet without any single serious competitor. As Monat also has suggested in discussing the community library:

> the idea of the community library as an active information center and not merely a repository for books, a center designed to serve a broad range of interests and diverse local institutions, has seldom been discussed much less explored operationally. A great information void exists in most medium-sized cities. Public agencies lack readily accessible and relevant information. There is no easy source of information for local merchants, financial institutions, and industries. And there is little done to publicize and exchange the information that is available within the community or region. The local library and district center possess the potential and already occupy the publicly accepted role that would support their development of this kind of information system.[8]

Nor do we see why public libraries should yield to the school interests with regard to children's services. Indeed, this goes against the trend of the times, which is for other agencies to step in where the schools have failed. In preference to abandoning their historical out-of-school educational role, public libraries might do better to demand the resources needed to devise adequate services for this constituency.

What does librarianship have to offer to a future information pro-

fession? We believe that in many ways librarianship is further along the road of professionalization than the newer information specialties. In the ideal of intellectual freedom, focused and infused with an advocacy commitment, may lie the rationale for the social purpose of the information profession. We could also argue that it exists, that it has been institutionalized and given legal and other mandates and guarantees of support, and that it has educational programs which have passed the inspection of and won acceptance by universities and accreditation bodies. But these institutional characteristics become an asset only if the profession is prepared to completely shift its institutional commitments to the extent of reallocating its resources and revolutionizing its practices. The support of a national program of library automation should be enlisted to effect these changes in libraries.

Manpower Factors

We are assuming that readers in the library profession share the conviction that the future projected for libraries at the beginning of this article is an undesirable one, not only because they are in and of librarianship, but also because in libraries there is a potential resource to be exploited for the social good. We hope the reader understands that the pessimistic future has a positive purpose and that it will have failed if it makes the reader only discouraged or angry. Yet the question must still be asked, are the gains of the past and the changes of the present sufficient bases for anticipating a future for libraries quite different than that posited in our projection. And our thesis would be, no, not without quite active and major interventions into library affairs which are not yet underway, or even recognized as necessary, in the field. Without such interventions we see little hope of overcoming the dominating force of tradition.

Let us focus on one major ingredient, the potential influence of manpower development on the future of the field. Our concern is not to evolve manpower policy for the field, but rather to suggest what the key variables are, their inter-relationship, and their importance in permitting the adaptation of the conventional library. As an aid then to those who influence policy choices, the following are ideas and suggestions as to the needs and desirable directions for the field. These developments are supportive of national programs of library automation and in themselves are crucial to this field's advancement.

The ability of the library profession to change is inextricably tied to its ability to attract, educate, and utilize manpower in the cause of

change. The nature of the task requires efforts to move along many fronts simultaneously at a more rapid and more comprehensive pace than is presently conceived in most quarters of the profession.

Our scenario has suggested a number of key manpower elements. Automation, centralization, and technician programs in themselves will markedly influence the numbers, the types, and the locations of library workers in the future. In the public library sector we see the union as a potentially powerful ally in insuring that the extensive retraining programs required get underway and are implemented. Yet there is a danger here for the unions may want the security of assured jobs in new roles, while for the foreseeable future library organizations should be characterized by fluid and changing work patterns.

When we consider expanded roles for libraries and the broader issue of preparation for information work in all settings, then other educational requirements emerge. New curricula will have to be devised for workers who undertake the information processing roles, for those who are prepared to undertake much more sophisticated and active user service roles, particularly with information communities which are still largely unsophisticated and unaware of their information needs, and for managers who can adapt and renovate the conventional library and manage the complicated information enterprise of the future.

We only hinted at another manpower need—the need to fill what is becoming a serious leadership vacuum in this field. By this we mean far more than managerial competency. Rather, the need is for people who are prepared to give new meaning and significance to the library's role in the local community and for articulate spokesmen at the regional and national levels who can crystallize issues, engender support for needed changes, and insure the position of libraries in future information networks. The shift required may be away from institutional leadership to professional leadership.

What should characterize this emerging leadership? It may need for its base a measure of activism, involving a disenchantment with the traditional, but also a zest for change and an ability to identify the key issues and platforms for change. This new leadership probably cannot hope to enjoy the following of the majority in the profession now. They will continue to lean toward the present institutional leadership whose values they share. Therefore, not only the proposals of the new leadership, but its political strategies as well, will need to vary from those of the past.

Clearly this field also needs to attract quite different types of people than have traditionally found their way into librarianship—more talented and capable people from the "hard" sciences, more who are behaviorally oriented, more men, more Blacks. Without this shift we see little prospect of changing the institutional stance of libraries from a passive to an active one.

In making the massive recruitment effort which we believe is called for, librarianship can no longer depend on the traditional "desire to serve." The profession through its associations and its schools must take a more aggressive position on salaries, so that the salaries in this field come to be not what a single person can live on, but rather what a married man with three children requires. It would be naive to proceed with recruitment under any other terms. And we must accept the fact that these new entrants expect to get ahead and move faster than has traditionally been the case in this field.

Recent recruitment efforts at Maryland offer some additional clues as to increasing the occupational attractiveness of the field. We have found that the newer information developments have reached the consciousness of students and are attractive to them. Another appeal being used, particularly with Black people, is that librarianship is a changing profession which welcomes people who would like to increase the social usefulness of libraries.

But improving economic incentives and rewards and shifting the appeal are only the beginning. Inevitably these new "types" will manage to shift the organizational environment of libraries so they will be supportive of wholly new relationships with clients and with regard to utilizing the computer. But the change process would be vastly speeded up and the recruitment task made easier if libraries could prepare now for their accommodation to new elements and come to deal frankly and openly with the internal consequences of change. External pressures for change may cause even the most conservative libraries to accept new people with differing talents and points of view. But with the demand for workers in newer information facilities, it is difficult to see how libraries can maintain their status quo commitments and at the same time provide job opportunities attractive enough to compete in the talent market of the information field. Over time we can hope the professional struggle will come to center around improving the organizational environment for professional practice rather than security concerns. Similarly, unless the educational programs genuinely respond to people with perspectives

and orientations quite different from the traditional humanistically, middle-class orientation, their efforts will not ultimately succeed.

Not even under the best of circumstances can librarianship place all its hopes on, or simply wait for, the new entrants into the field. There are also other manpower resources to be tapped. Indeed the strength of the information movement came from its ability to attract mature people from many walks of life who brought with them the sophistication and expertise needed to solve information problems. This involvement with those from the "hard" sciences is to some extent taking place now.

But there are other equally important involvements which are currently almost totally lacking. The issue of social responsibility requires insight from those in such fields as journalism, political science and sociology who are also concerned with problems caused by information imbalance. The search for viable positions on the information needs of one or another constituency and the legislative and other programs of action which should follow require linkups, interactions, and joint efforts with these related interests. Much of the ferment and innovation emanating from the field of education is also of direct relevance to librarianship. Here, too, librarianship should seek involvements with this field's innovators and spokesmen for change. Assistance in coming to grips with the essential question of relevancy of libraries may also come from new interaction with the library's users—and non-users—provided ways can be found to present the library's potential to them in new terms. As libraries shift their concern into their communities and identify with the needs of their constituencies, we can expect increased insight into how to make the library's role more meaningful.

The purpose of these reflections on manpower direction has been to emphasize the need for change and to suggest what may be the magnitude and directions of change.

Clearly the institution which will increasingly invite recrimination is the library school. Here, in particular, gradual change in the form of accommodation to newer information interests is woefully insufficient. From the schools, too, should come help in inserting into the professional ethos a sense of social responsibility which can be translated into ability to play socially responsive work roles. In this way the library school would become the dominating force in establishing the social purpose and utility of libraries.

Many of us in library education are becoming increasingly aware that we cannot much longer take people from all walks of education,

but primarily still the humanities, and hope to prepare them for future information roles in a one-year period. Neither can present programs much longer be justified on the grounds that we must fill jobs in libraries. The profession is becoming increasingly unwilling to let the schools keep their monopoly on entry into professional practice as demonstrated by their acceptance of those with less than full professional preparation for professional work roles. In government and industry there is almost no pretense that the library school path matters at all.

There is need for leadership everywhere in this profession; the way is open as never before for the library schools to play a leader rather than follower role in relation to the profession. Unrest is rampant; it is ideas and solutions to problems which are lacking. That the schools are not under pressure now to respond more dramatically means simply that this has not been the history of education in this field. Typically, the schools have followed rather than led practice. Without this expectation from the profession, the schools face the same task libraries face, which is to alter their image of themselves so that they respond in terms not of their history but of their potential and to the mandate implicit in their chartering.

It is not so much a question of whether the schools should risk the loss in enrollments if they lengthen their program, whether they will indeed be able to attract new types of students, or whether they have the faculty talents to implement new programs. Rather it is that they cannot afford not to make the effort. These problems must be viewed as challenges to be met, otherwise they will become mere excuses. The argument that the library field still wants the schools to prepare people for traditional types of positions is becoming a rationalization for maintaining the status quo.

The insertion of new faculties with new points of view should help the schools to escape the straitjacket of conventional ways of viewing libraries and may give them the courage to embark on new endeavors. Perhaps the most challenging educational aspect lies in the public and academic library spheres. A group in the Maryland School of Library and Information Services has recently been focusing efforts in these spheres. The approach to curriculum reform has been to determine work roles of the future and then to translate these into educational preparation terms. To give fresh perspectives and possible prototypes, various directions have been considered including the possibility of transplanting models from industry and government into

the more conventional library settings and the possible utilization of such "far out" activist efforts as the free press and switching center services. Programs to prepare citizen information specialists for work with the urban poor and to prepare undergraduate information specialists especially for work with today's student also are under consideration.

Efforts of this group suggest that students begin with or acquire a stronger base in quantitative areas and in the behavioral sciences than has traditionally been the case. This program would have both a strong theory base and, in some form, practical experience in the new work role so that the student would gain the preparation he needs while in school, and feedback would be received into the educational program. Work by this group is far from crystallized and is described here to indicate something of what could be the possible excitement and promise of the field today. Library education, in addition to responding to the profession's need for change, should at last put its programs on a par with the other professions, an important factor in its ability to compete for the talented people who will enter one or another of the professions.

In 1980, when librarians look back on this era, we believe it will have proved to be one of the most decisive—and hopefully exciting—in library history. To understand what happened they will have to view the library schools' response—or failure to respond—to librarianship's search for mature purpose and to the broader mandate of information access in this society.

References

1. Merton, Robert K. *Social Theory and Social Structure.* Rev. ed. Glencoe, Ill., Free Press, 1957, p. 568.

2. *Journal of Library Automation.* Chicago, Illinois. Vol. 1- . May 1968- .

3. King, Gilbert W., *et al. Automation and the Library of Congress.* Washington, D.C., Library of Congress, 1963.

4. White, Herbert S. "Toward Professionalism," *Special Libraries,* 60:73, Feb. 1969.

5. Leimkuhler, Ferdinand F., and Neville, Anthony E. "The Uncertain Future of the Library," *Wilson Library Bulletin,* 43:37, Sept. 1968.

6. Quoted in: Lanham, Richard. "Marian the Technologist?" *Computer Digest,* 4:10, Feb. 1969.

7. Taylor, Robert S. "Toward the Design of a College Library for the Seventies," *Wilson Library Bulletin,* 43:46, Sept. 1968.

8. Monat, William R. "The Community Library: Its Search for a Vital Purpose," *ALA Bulletin,* 61:1310, Dec. 1967.

ADDITIONAL REFERENCES

Becker, Joseph. "Information Network Prospects in the United States," *Library Trends,* 17:306-17, Jan. 1969.

Berelson, Bernard R. and Asheim, Lester E. *The Library's Public; A Report of the Public Library Inquiry.* New York, Columbia University Press, 1949.

Bundy, Mary Lee. "Automation as Innovation." *In* Paul Wasserman and Mary Lee Bundy, eds., *Reader in Library Administration* (Reader Series in Library and Information Science). Washington, D.C., Microcard Editions, 1968, pp. 369-76.

Bundy, Mary Lee. "Factors Influencing Public Library Use," *Wilson Library Bulletin,* 42:371-82, Dec. 1967.

Clapp, Verner W. *The Future of the Research Library* (Phineas L. Windsor Series in Librarianship, No. 8). Urbana, University of Illinois Press, 1964.

"Toward the Year 2000: Work in Progress," *Daedalus,* Vol. 96, Summer 1967.

Danton, Emily Miller, ed. *The Library of Tomorrow; A Symposium.* Chicago, ALA, 1939.

Eurich, Alvin C., and the Staff of the Academy for Educational Development, eds. *Campus 1980: The Shape of the Future in American Higher Education.* New York, Delacorte Press, 1968.

Gardner, John W. *Self-Renewal; The Individual and the Innovative Society.* New York, Harper & Row, 1964.

Hirsch, Werner Z., *et al. Inventing Education for the Future.* San Francisco, Chandler Publishing Company, 1967.

Jouvenel, Bertrand de. *The Art of Conjecture.* Translated from the French by Nikita Lary. New York, Basic Books, Inc., 1967.

Kahn, Herman, and Wiener, Anthony J. *The Year 2000: A Framework for Speculation on the Next Thirty-Three Years.* New York, Macmillan Company, 1967.

Licklider, J.C.R. *Libraries of the Future.* Cambridge, Mass., M.I.T. Press, 1965.

McDermott, John. "Knowledge is Power," *Nation,* 208:458-62, April 14, 1969.

Overhage, Carl F. J., and Harman, R., eds. *INTREX: Report of a Planning Conference on Information Transfer Experiments.* Cambridge, Massachusetts Institute of Technology, 1965.

U.S. National Advisory Commission on Libraries. *Library Services for the Nation's Needs: Toward Fulfillment of a National Policy—Report of the National Advisory Commission on Libraries, July 1968. In* U.S. *Congressional Record,* 2d Sess., 1968, E9355-68. See also the special studies on which the report is based, as listed on pp. E9366-67; and 114:E9355-E9368, Oct. 14, 1968.

U.S. President's Science Advisory Committee. *Science, Government, and Information: The Responsibilities of the Technical Community and the Government in the Transfer of Information; A Report.* Washington, D.C., U.S.G.P.O., 1963.

Vosper, Robert. "The Computer: No simple Cure [Reply to F. F. Leimkuhler and A. E. Neville]," *Wilson Library Bulletin,* 43:43, Sept. 1968.

Wasserman, Paul. *The Librarian and the Machine: Observations on the Applications of Machines in Administration of College and University Libraries.* Detroit, Gale Research Company, 1965.

Wasserman, Paul. "The Library and Information Professions in a Time of Change," *Pacific Northwest Library Association Quarterly,* 31:134-45, Jan. 1967.

Ways, Max. "The Era of Radical Change," *Fortune,* 69:112-15+, May 1964.

Bibliographic and Technical Problems in Implementing a National Library Network*

HENRIETTE D. AVRAM

THE PROBLEMS FACING the planners of automated library networks are rooted in the complexities of organizing and managing a vast flow of bibliographic information and its interface with users. Telecommunication equipment transmitting data in the form of electric signals, electronic memories holding large stores of information, and computers manipulating the data and graphic displays for human interaction are technological means for performing network functions more effectively than has been possible in the past. They do not in themselves, however, make networks possible.

Becker has listed the following among the problems and obstacles to be overcome: the development of acceptable criteria for determining what is to be placed on the network, clarification of the roles of network participation, agreement on network organization and operation, and the investigation of its social, legal, financial, and technical implications.[1] The designers and operators of a network must understand the need to cooperate and to compromise in determining objectively and rigorously which areas are most susceptible to cooperative action and which will have the greatest benefit in the shortest possible time. At the very least, library networks require common languages and common procedural conventions.

The underlying concept of library networks is well-established. For many years, libraries have been cooperating to make the greatest

Henriette D. Avram is Assistant Coordinator of Information Systems, Library of Congress.

* The author wishes to acknowledge the major contribution of John C. Rather to the development of the concepts contained in this paper. Our dialog over many hours provided a powerful catalyst in crystallizing my thinking about a down-to-earth presentation of problems in the construction of a national library network. I also appreciate his work in editing the paper for publication.

possible use of available information resources by sharing them through arrangements of varying degrees of formality. In a comprehensive survey of the development of library cooperation, Esterquest mentioned twelve types, including interlibrary loan, union catalogs and lists, regional bibliographic centers, cooperative storage, cooperative acquisitions, and cooperative cataloging.[2] Each of these efforts has been regarded, in its time, as providing the solution to a pressing problem. Nevertheless, the difficulties in making such cooperative relationships work effectively have led to the downgrading of some and to the abandonment of others.

A major impediment to the success of library cooperation has been the difficulty of maintaining a regular flow of up-to-date bibliographical information among libraries. In the last several years, however, two developments have given hope that this situation might be alleviated. The expanded operations of the Library of Congress under the National Program for Acquisitions and Cataloging have helped to speed the production and distribution of catalog records for current publications. In the technical domain, the MARC Pilot Project and the MARC Distribution Service have demonstrated the feasibility of distributing catalog data in machine-readable form.

The second of these developments has taken place in a climate highly favorable to automation in libraries all over the United States. The possible applications of the computer to library operations are being explored in the belief that their efficiency will be enhanced. Using the new technology, libraries should be able to attain greater speed and flexibility in creating, updating, and disseminating bibliographic information. The anticipated success of this effort has, understandably, rekindled enthusiasm for sharing resources through immediate access to a common bibliographic store and rapid transfer of information within a network of libraries.

Experience has already shown, however, that even automating individual libraries requires solutions to many difficult problems. Establishing a workable automated library network involves difficulties of still greater magnitude. This paper takes a broad view of problems involved so as to identify and to analyze basic issues that tend to be glossed over in our eagerness to approach the goal. Inevitably, the discussion may overlap topics covered in other papers in this issue. The parts, functions, and attributes of a library network are so inextricably related that the examination of any element must impinge on other elements.

Bibliographic and Technical Problems

The Concept of "Library Network"

The literature in the past few years is rich in discussions of future international networks, national networks, state networks, regional networks, etc. Many network plans have been put forward. Nevertheless, the lack of a generally accepted definition of a library network causes confusion. Becker and Olsen defined a network as "an interconnection of things, systems, or organizations. Adding the adjective *information* to network allows the concept to be defined with greater precision. In an information network, more than two participants are engaged in a common pattern of information exchange through communications for some functional purpose." [3] Within this definition, the authors described the ideal information network as exhibiting the following characteristics: formal organization, a communications system, bi-directional operation, a directory look-up system to identify the unit that must be able to respond to a query, and a switching capability to determine optimum routes.

On the basis of this definition, a single library can be shown to be an information network for its staff and users. The library has a formal organization governed by established policies and procedures. The staff is grouped into divisions with distinct functions (e.g., cataloging, reference, circulation). The interfaces among divisions through individual staff members using common files and the interaction of staff and files with users constitute the communication system. The directory look-up is provided by the bibliographic control apparatus which comprises all of the files for locating items in the library collection. The main catalog affords the most complete coverage; some of the other files are tangential to it; the contents of others overlap.

The library staff serves the function of a switching mechanism to determine optimum routes for queries. For example, a user seeking an unbound serial will be referred to the serial record instead of the main catalog. The flow of information is bi-directional. A reference librarian responding to a user's query uses one or more of the bibliographic control devices. If he finds them inadequate or incorrect, he notifies the cataloging division (in effect, a cataloger) which may change or add to the information in the files.

The requirements for the ideal "single library" network include accurate and up-to-date information. The network communications function efficiently when they provide an "immediate" answer to a query even if the answer is negative. (The word "immediate" in this context means within the required response time, whatever that may

be.) A negative answer is hardly satisfactory, however, particularly if it results from a failure to generate information rapidly enough (as might be the case when there is a cataloging arrearage). An even more common difficulty arises when the information is somewhere in the network but technical or organizational shortcomings inhibit its flow to the desired point. This situation occurs frequently when the bibliographic control apparatus comprises many separate files.

This concept of a "single library" network can be projected into a national network of libraries, using virtually the same framework with additional hierarchies or levels, and increased bi-directional capability. The switching mechanism no longer depends on individuals but rather on well-defined nodes or centers that transmit requests to the appropriate information resource by the most expeditious route and transmit the relevant material back to the source of the query.

If a network were organized with major regional centers as the intermediate nodes, it might seem that to avoid traffic congestion when those centers access the national bibliographic store, the entire data base would have to be entirely duplicated in several places. This could be avoided if each intermediate node assumed national responsibility for a subset of the total information base; that is, records that could be specifically categorized by language, date, subject, or type. A hypothetical network for sharing cataloging data might have regional nodes that maintained union catalogs for their respective areas and also served as distribution centers for particular segments of the national data base. The national center would distribute its output to every regional node which would keep all records for a prescribed period. As records were supplied to libraries in a region, they would be posted to the regional union catalog and to the national bibliographic store. At the end of the stated retention period (say, a year), the regional node would delete all records outside of its national responsibility if they had not been added in the regional catalog. This procedure would reduce the file in the regional nodes. The rapid flow of information in the network would be facilitated by the ability of a regional node to satisfy many requests from its member libraries and to route others to another regional node known to be responsible for a given category of record. In addition to being the primary distribution point for newly generated records and possibly a regional center in its own area, the national bibliographic store would serve as the court-of-last-resort for requests outside the scope of any regional center.

Bibliographic and Technical Problems

As long as the network designer is using paper and pencil and hypothesizing without constraints of cost, organization, legal implications, etc., there is practically no limit to the kinds of networks that can be created by assigning different values to the building blocks and assembling them in different ways. In practice, however, the success of a library network will depend on the extent to which it satisfies certain basic requirements.

Requirements for a Library Network

The preceding section was not intended to provide a blueprint for a library network. Its main purpose was to show that many of the problems of a library network and those of an individual library differ in degree but not in kind. Thus, it is not surprising that difficulties that have plagued libraries for decades still persist in the age of the computer. In this section, some of these problems will be explored in more detail in two main categories: 1) standardization of the bibliographic record and 2) technical considerations.

A third element which is not considered in this paper is the need for sufficient information in the network. In absolute terms, the information store depends on the level of financial support of the libraries comprised by the network. If they are unable to acquire and catalog the materials to satisfy the needs of their users, the amount of information available will be below the required level. In relative terms, however, the solution lies in the adequacy of the links among the libraries. The dominant purpose of standardization of the bibliographic record and the provision of means of rapid communication is to facilitate the pooling of bibliographic information. If these conditions are met, the amount of information in the network will tend to equal the sum of the information in all of the individual libraries.

Standardization of the Bibliographic Record. It is easy to minimize the difficulties in creating a bibliographic record that is standardized in format and content. To achieve the ideal result, agreement must be reached on four major points.

1) There should be a standard set of rules for describing and analyzing bibliographic items. Great progress has been made toward this goal by the formulation and general adoption of the Anglo-American Cataloging Rules [4] for the creation of catalog headings and the description of bibliographic items. The situation with respect to subject analysis is somewhat less clear. Library of Congress subject headings,

LC classification schedules, and the Dewey Decimal Classification are widely used in American libraries. In practice, however, the application of classification schedules and subject heading lists may differ from library to library (or even cataloger to cataloger) because there are no generally accepted procedures for analyzing the materials being cataloged. The development of clear guidelines for subject analysis is highly desirable, particularly if local records are to be posted to a national data base. The task is formidable, but any success in accomplishing it will contribute significantly to the consistency and manageability of the data base.

It is obvious that equally difficult problems must be solved before international standardization is possible. Nevertheless, it is encouraging to note the International Meeting of Cataloging Experts held in Copenhagen in August 1969. The purposes of the meeting were to review cataloging developments and to examine the prospects of cataloging advances through standardization and mechanization and, in this connection, to consider the national bibliographies, the Shared Cataloging Program, and the production of cataloging data in machine-readable form. The aim was to arrive at conclusions of practical value which will further international uniformity in cataloging. It is to be hoped that this meeting has set the stage for further progress in international cooperation and provided the climate for the advancement toward the ultimate goal of a true "sharing" of information resources.

2) Bibliographic records should be prepared in relation to a standard data base. The principal aims of descriptive cataloging as commonly practiced by libraries are: a) to provide a unique description of each item, b) to bring together the works of an author, and c) to bring together editions of a work. The first point can usually be resolved by rules alone. The second and third, however, require the establishment of a consistent form of name for a particular catalog and the coordination of each new record with existing records. It is obvious that both a file of name authority records and a known data base of bibliographic entries must be readily accessible to the originators of catalog records to insure consistency. Without a method for distinguishing items as being unique or for relating them to other items in the file, bibliographic control becomes a Tower of Babel for the librarian and, in turn, the user.

The problem becomes evident when the experience of the National Union Catalog (NUC) is examined. Reports for the same biblio-

graphic item are frequently received in widely different forms. In a discussion of the NUC, a recent study concerned with the conversion of retrospective catalog records to machine-readable form states "that wide variations in bibliographic description would make it difficult to identify many of these records as being for the same item." [5]

Some of the confusion among catalog records is attributable to differences of interpretation of the rules and information on the publication. Much of it, however, occurs because outside libraries cannot conveniently obtain up-to-date information about the preferred form of heading. A search of the book catalogs of the Library of Congress is often time-consuming and sometimes yields a heading in an obsolete form. Ready accessibility to a current source of established names is one of the benefits that an automated library network should provide.

The records being disseminated by the MARC Distribution Service constitute an acceptable body of standardized bibliographic data for English language monographs. It is not yet a complete source of cataloging data for subscribing libraries, however, because it does not provide information about *see* and *see also* references for the headings used in the records. The Library of Congress is aware of this drawback and plans to distribute reference control information as soon as possible. When this is done, the MARC data will become a subset of a true national bibliographic store.

The body of standard cataloging data in machine-readable form will be enriched as retrospective records are converted. The RECON Pilot Project now under way at the Library of Congress promises immediate benefits through conversion of approximately 69,000 English-language records. Expansion of the MARC Distribution Service to cover other languages and a large-scale conversion project for retrospective records are other possibilities in the foreseeable future.

3) There should be a standard set of rules for structuring machine-readable records for all forms of material and labeling their data elements. There is a growing acceptance of a basic structure for a format that prescribes the physical layout, leader, directory, control fields, and variable fields. This structure was designed in collaboration with many groups and is being considered as a national standard by the United States of America Standards Institute upon the recommendation of its Section Committee Z39 (Library Work, Documentation, and Related Publishing Practices). The format has been adopted by the American Library Association, the Special Libraries Association, the National Libraries Task Force on Automation and Other Coopera-

tive Services, the Association of Research Libraries, the Committee on Scientific and Technical Information (COSATI), the Federal Library Committee, and the British National Bibliography. The MARC II format used by the Library of Congress conforms to this proposed standard.

Progress is also being made in the definition of content designators that explicitly identify data elements for different forms of material. Both the Library of Congress and the British National Bibliography are using essentially the same format for bibliographic descriptions of monographs. The Library of Congress has issued a recommended serials format to elicit comments from the library community. The Library of Congress has also designed and is using a format for single-sheet maps and is making progress toward definition of content designators for audio-visual material. COSATI has defined content designators for technical reports to be used by executive agencies in the federal government. All of these formats have the same basic structure.

Agreement on a common format has made possible the exchange of machine-readable bibliographic records between the Library of Congress and the British National Bibliography. The potential advantages have led Coward to assert that "a MARC record service must transcend national boundaries; it must have an authority which makes its records acceptable to librarians anywhere in the world; and it must strive to be as complete as is humanly possible. I do not think that there is any future in attempting to produce a national service unrelated to other national services." [6] The growing international interest in the transmission of bibliographic data in this form is also exemplified by a French translation of *The MARC II Format.*[7]

4) There should be a standard degree of completeness of the data elements in a machine-readable record. Within the basic structure, records can vary in two respects: content designators can be simplified and data elements can be omitted. For example, a name entry could be identified simply as a name rather than defined by type, or, the bibliographic description could be streamlined by omitting notes. The Library of Congress and the British National Bibliography have taken the position that the records they distribute should be as rich in detail as possible. Their premise is that, on the basis of present knowledge, it is impossible to define rigorously every potential use of a machine record. Therefore, the difficulty and cost of augmenting a record make it prudent to provide a full record even if unwanted

items may be deleted later. On the national level, it seems unwise to do less than is now being done.

With more study of the effect of different levels of content designation in MARC records, it may be possible to simplify encoding bibliographic records without detriment to a cohesive library network. The minimum degree of completeness of bibliographic data will be determined by what is required for uniqueness in the master data base (see item 2 above) if there is to be a bi-directional flow of data. Where data flows only in one direction (for example, from regional center to a local library), records at the lowest echelon in the network (local library) may be less complete than those at the national level. Levels of MARC records and their implications will be studied during the RECON Pilot Project.

Technical Considerations. Many problems must be solved and questions answered before a true national library network can be created. In a recent review, Bregzis stated:

> Although the concept of a central bibliographic data file has lately become quite popular in library automation plans, the problems arising from the massive size of such a file, the complex logical structure of records, multidimensional interrelationships among records, and technical constraints associated with data storage, access, and telecommunication have generally been overlooked or dismissed. . . . Whether library technical processing, as it is presently known, permits large-scale consolidation even under computer control is an open question and not beyond doubt.[8]

Although it is impossible to discuss every technical consideration vital to the creation of a successful national network, this section will mention some important factors that are often ignored by network planners.

1) The dynamic characteristics of bibliographic records are frequently underestimated and as a result not enough weight is given to the requirements for updating. Not only does the increased growth rate of printed material affect the problems of the initial control of library holdings, but maintenance of the bibliographic records themselves becomes increasingly difficult, albeit indispensable. A recent analysis of the extent of changes in Library of Congress catalog cards in the RECON study[9] provides convincing evidence that ignoring changes made to records in an active catalog would result in a significant loss of the quality of the cataloging information.

The MARC Distribution Service does reflect most of these changes. Any substantial change in the LC Official Catalog (e.g., a change in main, added or subject entry) triggers a revision of the machine-readable record and this updated record is, in turn, distributed to the subscribers as a "change" record. However, some "housekeeping" changes are not distributed to the user. For example, the Library does not update the MARC record of an incomplete set as volumes are added unless the record contains a content note; the change is made only when the set is complete.

This experience stimulates many questions in the context of a formal network. What is involved in maintaining bibliographic quality and accurately reflecting holding information? Would the records of every node in the network be updated? If every "change" record distributed by the national center has to be inspected by regional nodes to determine if the original record is in the regional system, there is a cascading effect throughout the network structure. What are the cost implications of the additional flow of information through the system? Would failure to update at every level of the network result in the problems of inconsistency that exist today?

Assuming files will be organized in the same fashion as they are today (i.e., an in-process file, and a catalog record file), regional nodes will be required to store MARC records in a separate file until they are required in the regional system at least for some period. Any change record received might have to be compared against all files since at any point in time, there is no way to know where in the system the record resides. An alternative scheme could be the maintenance of an index by LC card number which would have an associated communications field to indicate in which file the record is presently located. This index would require updating in its own right as the record moved from one status to another. The problem is compounded many times when one begins to envision the maintenance of authority files, the required links back to the bibliographic records, and the complicated machine procedures required to implement what really can be considered a "network" in its own right.

2) The method of data organization for the storage, retrieval, and maintenance of machine-readable files is heavily dependent on the requirements of the users of the system, e.g., optimum retrieval capability must sometimes be sacrificed to achieve a balance between maintenance and search efficiency. The size of the files and the computer hardware and storage media available are other variables that must be considered in system design.

Bibliographic and Technical Problems

The question of how best to structure information is not unique to the planners of automated library systems. A number of techniques have evolved, e.g., direct, random, inverted, indexed sequential, multilist, ring, tree, etc. Depending on the scheme chosen, there are associated problems such as space management, dynamic storage allocation, nesting, paging, address calculation, etc. It is not the purpose of this section to give a detailed presentation on file organization, but rather to emphasize the need to understand the complexity of data organization when library networks are being considered. The interested reader will find more detailed exposition in several excellent articles (See Additional References).

Assuming the existence of a suitable computer and storage media, the designer must ask himself early in the planning stages, "What are the elements in bibliographic data that are frequently used as search arguments and therefore should be selected as keys?" Beyond the commonly accepted elements of author, title, and subject (and not necessarily in that order), there is little agreement among librarians as to a rank order of importance of other descriptive items. And, so far, catalog-use studies have failed to provide this type of substantive data. Even if this information were available it would probably not be possible to design a system that would provide 100 percent satisfaction for all users. To do this would require making available every data element as a key. If the system to be designed was an inverted list structure, the use of every data element as a key would require a dictionary of attributes and machine addresses which might be as large or larger than the data file itself. Although this technique satisfies the retrieval requirement, the problem of maintenance of a large dictionary is difficult and costly. Other file organization strategies could be employed, but without sufficient knowledge of the most useful access points, there is little basis for evaluation of one technique over another and planning cannot be accomplished with confidence.

Another important question is whether the file structure for library networks will be the same for all forms of material. For example, posting holdings information for serials is a problem of far greater magnitude than the related task for monographs because serials frequently change titles, merge, and undergo other transformations. The very nature of serials appears to demand an organization of the files that will allow linking of one record to another so that regardless of the title or issuing body requested, a query will be satisfied or a new item posted.

How will we organize the subject heading file and the name reference file and guarantee that any addition, deletion, or change to one of these files will manifest itself in the main bibliographic file when applicable? What is the complex organization that will provide linkage from element to element and record to record within a file, and element to element and record to record from one file to another or perhaps to several?

If the National Union Catalog of 7.5 million records is considered as an approximation of the size of the bibliographic data base at the national node and each record is estimated to be 500 characters in length, the storage capacity at the national level would have to accommodate 3.75 billion characters. Added to this already voluminous count would be name reference records and the subject heading records plus the characters required to provide the linkage or the overhead of the system. Needless to say, the technique chosen for the organization of bibliographic files must undergo a careful evaluation of cost of overhead versus the advantages of potential retrieval.

Even if the decision were made to plan library networks based on current cataloging only, the problem, though not quite as formidable to start, would still fall under the heading of "large files" with the same complex relationships and would present the same perplexing technical considerations to the designers. (The Library of Congress catalogs approximately 200,000 new titles per year and it is estimated that the growth rate is 5 percent.)

Because of the great size of the files, one is forced to question whether it is necessary to store the entire bibliographic record in digital form or whether it would suffice to store a select number of data elements in digital form with linkage provided to the record in a slower microimage storage (this concept was suggested to the author by Allen Veaner in a different context than this article). This is still another facet of the problem that is deserving of careful evaluation, measuring need against cost.

3) A related but unsolved problem is that of the composition of a search code. A search code is a string of characters made up of selected characters from one or several data elements, e.g., author, title, imprint, date, etc. A search code may serve several purposes: a) to shorten the character comparison required between the search argument and the records in the file (assuming the search code is a key associated with each record and the same algorithm used to develop the search code key is used to develop the search code for the search

argument), b) to calculate an address in computer memory where pointers exist to the bibliographic records that contain the data elements from which the search code was derived, c) to cope with name variants caused by misspelling, transcription errors, name on title page not identical with established form of name, change of name due to inversion of the name, etc., and d) to cope with title variants caused by misspelling, transcription errors, lack of definite knowledge of the title, etc.

Some interesting work is already in progress in this area.[10] It might well behoove the investigators in personal name searching in the library community to explore what already has been accomplished elsewhere. The problems of searching files where a principal access point is by personal name is not limited to bibliographic work. A great deal of effort and funds have been expended to develop methods to search other large machine-readable files. Insurance companies, airlines, and the Social Security Administration face this problem and the problem of variant names is a severe one in the Immigration and Naturalization Service, the Federal Bureau of Investigation and, in fact, all intelligence activities.

Assuming storage capacity, some optimal file organization, the development of search codes, etc., does the operation of a library network imply that the same hardware configuration with the same software is implemented at all major levels in the network? What is the price tag to fulfill a national responsibility?

4) If networks are to become a reality, the coded representation of symbols required for the encoding, storage, retrieval, display and transmission of bibliographic records must be standardized. A review of the work performed during the last decade indicates not only the awareness of the need for standardization of codes for the representation of characters, but also the need of the associated hardware devices to input and to display these characters.

Progress in the development of data input devices has generally lagged behind the technical development of other functions in automated systems. Therefore, data input tends to remain the slowest and least efficient function and the one most prone to error. This condition is exacerbated by non-numeric processing where the data may contain virtually any symbol.

The Library of Congress designed a character set for the representation of Roman-alphabet languages and it includes many special characters and diacritical marks for those languages. The MARC Dis-

tribution Service utilizes this character set. It was recognized that no print train/chain or terminal device existed that could display this character set at present. The philosophy at LC was not to design tomorrow's systems constrained by today's technology. Therefore, for some time to come, the recipient of MARC tapes will have the burden of translating the MARC character set into the character set available to him at his installation. This may mean printing substitute characters for nonprintable ones, or stripping certain characters out completely. The result is extra processing time for each user and in addition, tailor-made software for each user depending on his particular display device.

In a library network, the need for a standard character set is even greater. If nodes in the networks use different sets of coded representations, many tailor-made computer programs would be needed at each node to translate the data received from other nodes.

The non-Roman alphabet languages pose problems of additional complexity. If a unique code on an input device signals an escape into another alphabet (e.g., code plus C equals Cyrillic alphabet), it is possible to use the standard keyboard to encode many alphabets. However, how does one display the language that has been recorded? Again, with the exception of photocomposition devices which most libraries cannot afford, the available devices are not capable of displaying the diversity of characters required for many different alphabets.

The design of hardware that will satisfy the requirements for the efficient and economic input and display of bibliographic records is still in the developmental stages. The technology must be constantly monitored and systems modified as the state of the art improves.

5) Large data banks must reside at more than one installation if the system is to be capable of satisfying a user even if one of the nodes is not in operation. It is not feasible to consider the transmission of large files on demand from one node to another in the eventuality of down-time on the prime system, i.e., the node the user should be accessing. Therefore, to maintain back-up, procedures would have to include the storage of files at several centers in the network. What is the impact of the duplication of the machine-readable library files to guarantee back-up? Will it be necessary to provide back-up at all echelons of a library network, i.e., the national data store, the regional data stores, etc.?

6) The installation of a data transmission network implies the

linking of many pieces of hardware. When transmission problems occur, it may be difficult to determine which device is at fault—the computer, the transmission equipment, the data sets, etc. This problem has a tendency to decrease as the network expands since alternate paths can be used to isolate the cause. In most instances diagnostics do exist that will determine computer failure. However, valuable computer processing time is lost during diagnostic processing and transmission time during alternate path testing. The maintenance of the hardware of the network must be taken into consideration when deriving cost factors and projecting optimum utilization.

7) The monitoring of centralized data banks in a library network will require that users have unique identification numbers for accounting purposes as well as to insure that files cannot be altered by unauthorized individuals. Considerable research is going on in this area but the systems designed to date are limited and subject to being compromised. Some of the proposed techniques look promising but as yet most have not been implemented and tested. There also has been no evaluation of the cost of any of the proposed methods in terms of hardware and/or processing time. Although library files are not sensitive in the same sense as files such as those of personal dossiers, the integrity of bibliographic files could be damaged either by accident or unauthorized use. Therefore, a safeguard scheme must be an integral part of the operation of a library network utilizing centralized data banks.

Outlook

Given the reality of these bibliographical and technical problems, what is the outlook for library networks? Librarians who have been involved with automation know that the design of a system is but one phase of the implementation of automated procedures. There comes a time when generalization must cease and we must face the hard facts of "how to do it." Networks may be conceptually and technically feasible, but there is a long, difficult road to travel between here and there. The enumeration of problems in this article is not intended to lessen creative drive but to caution that separating the operable from the speculative is a necessary prerequisite for moving ahead. Those actively engaged in network planning may find consolation in Machiavelli's observation: "There is nothing more difficult to take in hand, more perilous to conduct, or more uncertain in its success, than to take the lead in the introduction of a new order of things."

References

1. Becker, Joseph. "Information Network Prospects in the United States," *Library Trends,* 17:306-17, Jan. 1969.

2. Esterquest, Ralph T. "Co-operation in Library Services," *Library Quarterly,* 31:71-89, Jan. 1961.

3. Becker, Joseph, and Olsen, Wallace C. "Information Networks." *In* Carlos A. Cuadra, ed., *Annual Review of Information Science and Technology.* Vol. 3. Chicago, Encyclopaedia Britannica, 1968, pp. 289-90.

4. *Anglo-American Cataloging Rules.* Prepared by the American Library Association, the Library of Congress, the Library Association, and the Canadian Library Association. Chicago, ALA, 1967.

5. *Conversion of Retrospective Catalog Records to Machine-Readable Form; A Study of the Feasibility of a National Bibliographic Service.* Prepared by the RECON Working Task Force, Henriette D. Avram, Chairman, John C. Rather, ed. Washington, D.C., Library of Congress, 1969, p. 109.

6. Coward, Richard E. "The United Kingdom MARC Record Service." *In* Nigel S. M. Cox and Michael W. Grose, eds., *Organization and Handling of Bibliographic Records by Computer.* Hamden, Archon Books, 1967.

7. U.S. Library of Congress. Information Systems Office. *Le format MARC II.* Translated by Marc Chauveinc. Grenoble, Bibliothéque universitaire de Grenoble, 1969.

8. Bregzis, R. "Review of *The Regional Library Center in the Mid-1970's; A Concept Paper* by Thomas Minder," *Computing Reviews,* 10:273, June 1969.

9. *Conversion of Retrospective Catalog Records . . ., op. cit.,* pp. 141-62.

10. Ruecking, Frederick H., Jr. "Bibliographic Retrieval from Bibliographic Input; The Hypothesis and Construction of a Test," *Journal of Library Automation,* 1:227-38, Dec. 1968.

ADDITIONAL REFERENCES

Dodd, George. "Elements of Data Management Systems," *Computing Surveys,* 1:117-33, June 1969.

Knuth, Donald E. "Information Structures." *In* Donald E. Knuth, ed., *The Art of Computer Programming.* Reading, Mass., Addison-Wesley, 1968.

Climenson W. Douglas. "File Organization and Search Techniques." *In* Carlos A. Cuadra, ed., *Annual Review of Information Science and Technology.* Vol. 1. New York, Wiley, 1966, pp. 107-36.

Hardware

J. FRANCIS REINTJES

THE LIBRARY COMMUNITY is examining with great interest the claims which technologists are making about the ability of their contrivances to alleviate current library problems. The rapid rate of build-up in recorded information makes it imperative to give recognition to any device or system that may help to reduce library operational costs, curtail capital investments, or assist in providing more effective library services. The relevant technologies center on the digital computer and its associated input/output equipment, on wideband and narrowband communication including wire and wireless communication, and on microphotography. At least in principle, exploitation of these technologies should enhance the effectiveness of the library as an information tranfer center because of their power as information transfer mechanisms and their ability to streamline library business management.

The effectiveness of any technological innovation which may be introduced into the information transfer side of library operations will ultimately be determined by the response of the library-user community to the innovation. The user-community yardstick of acceptability includes scales of cost effectiveness, information-retrieval effectiveness and personal convenience. Thus in the design of hardware for library services, full recognition of the characteristics and requirements of the consumer of the services must be given.

The issue of how a digital computer should be organized for the library environment is at present unresolved. The options are several. One may choose, for example, to employ separate machines for business operations and information retrieval. Alternatively, a single general purpose machine, capable of handling both functions can be specified. Each approach brings forth pro and con arguments. The dual machine configuration is obviously the more expensive, particu-

J. Francis Reintjes is Professor of Electrical Engineering, Massachusetts Institute of Technology, Cambridge.

larly if each machine remains idle for an appreciable part of each day. Nevertheless this approach may be necessary in order to avoid user dissatisfaction with the information-retrieval capabilities of the machine-oriented library system. Should use of the computer for business management induce prolonged delays in the execution of requests for information retrieval, the user community may easily become disenchanted with the whole computer-oriented operation. Nevertheless, a single computer for both information retrieval and for library business management is attractive if management operations can be performed at times which do not degrade the quality of service rendered to the library user.

New possibilities for computer architecture for library systems have been opened through the advent of large public utility-type computing machines which feature time-sharing of the facility among many users who may wish to engage it simultaneously and a capability which allows each user to carry on a dialog with the machine as he negotiates for stored information. In computer parlance this latter capability is dubbed *on-line interaction* between user and machine.

On-line interactive computer systems have exciting implications for information transfer. Inherent in these systems is the ability to engage them from terminals located remotely from the machine. The convenience of remote engagement directly by the person seeking information and his ability to negotiate back and forth with the machine in real-time as he zeroes in on precisely the piece of information being sought, make this type of computer configuration a most attractive item of equipment for the library of the future.

Let us therefore examine the on-line interactive machine in more detail. We are at present only on the threshold of the on-line interactive computing machine era. Although many such machines are in operation, they have been designed for a variety of purposes and with a broad range of performance capabilities; none have been tailored specifically to library applications. For information retrieval, the machines work this way: the data base, formatted in accordance with a plan developed by librarians and computer programmers, is located in a secondary storage device of the computer. Typically the data base consists of a catalog of information about the documents being held in a collection. (The problem of storing full text in a computer is discussed below.) The storage device may consist of a set of magnetic disks which looks very much like a stack of phonograph records, or other forms of magnetic storage such as magnetic drums or magnetic

cards. Also stored in the machine is a set of computer programs which serve to retrieve the stored information when the programs are used in accordance with prescribed rules. Since time-sharing computers can accommodate the computational or data processing requirements of a heterogeneous community of users, one may also expect to find stored in such a machine, in addition to the library data base and retrieval programs, a wide variety of other types of computer programs totally unrelated to the library function. Finally, there is also stored in the machine a set of executive, or control, programs which serves to manage the flow of information throughout the system.

In order to engage the time-sharing computer, the user operates one of two types of consoles: either a typewriter console which is very similar in physical appearance to an ordinary typewriter, or a cathode ray tube (CRT) console consisting of a typewriter keyboard for inputting information and a television-like cathode ray tube for displaying output information. Examples of these consoles are illustrated in Figures 1, 2 and 3. Great care is taken in the programming of these machines to make it easy for the user to express his commands, and an important feature of the programming is that it permits the user to interrupt his program and to alter his procedure as he progresses to-

Figure 1. Example of a Typewriter Console

Figure 2. Example of a Storage-Tube Display Terminal (*Courtesy Computer Displays Inc.*)

ward his end objective. Thus, the user, in effect, carries on a dialog with the machine and works along with it as he solves his problem in an *ad hoc* manner. The power of the on-line interactive feature is that it enables user behavior to be dynamic; a user is able to make decisions on the basis of partial results and to plan his strategy for information retrieval as he goes along. The power of the time-sharing feature is that many users can engage the machine simultaneously without serious degradation in machine performance. While one user is thinking about his next step, the machine is actively serving another.

Central to the time-sharing computer concept is the fact that while one user is pursuing his problem at his console, many other users are pursuing theirs at other consoles. Each user is served in turn for short periods of time; the frequency with which the machine returns to an individual user depends on the number of users engaging the machine and the complexity of their problems. Herein lies one of several unresolved issues with respect to these machines in the library appli-

cation. Granted that on-line interactive machines are ideal for information look-up purposes, it is still unclear at this point in time that they can be designed to provide prompt, efficient service for the total community served by a library. To date these machines are essentially untested in the true library environment, although some experimental results are available and experimentation is still underway. For the next four to six years, it appears the answer to this question will be influenced by the following interrelated factors:

1) the quantity of information that is stored in the computer. This factor is, in turn, related to the size of the operational library involved;
2) the size of the on-line user community, including the library user community and all others;
3) the manner in which the data base is formatted and the character of retrieval programs that are provided for accessing the data base; and
4) the hardware and software organization of the on-line time-sharing system itself.

At present it is possible to maintain 10^8 to 10^9 English words in storage, and this number is growing. Let us assume we wish to computer-store a catalog for a library containing 10^5 volumes, that we wish to devote 10^2 words to a description of each catalog entry, and that we wish to have both subject and author as access points. The total number of words in such a catalog is imprecise, since it depends upon the detail and depth of subject indexing. However, it is evident that a storage capacity of 10^8 words is ample, and that with careful file organization, space for growth will be available.

As a means of providing an improved catalog look-up system, one would like to think in terms of an expanded catalog in which individual articles of each issue of professional journals are indexed and indexed in depth. It is conceivable that a catalog of this type for our holdings of 10^5 volumes would demand a severalfold increase in storage capacity and thus would be taxing current storage technology to its upper limit. Storage capacity would then be an especially crucial matter if these computer facilities had to be shared with other users whose interests were unrelated to the library application.

Although most computers would be overtaxed under the above circumstances, the technology for building larger machines exists. For example, at the Atomic Energy Commission facility at Livermore, California, enough computer storage capacity is available to accom-

modate the alphanumeric content of approximately 10^5 books of 300 pages each. This storage capacity amounts to approximately 1.5×10^{11} bits.

The size and composition of the time-sharing user community is certainly a factor in the determination of the effectiveness of these machines in a library environment. Machines that can accommodate thirty to forty users simultaneously are now operating and others are being developed which will accommodate many more. However, as the user population on a machine increases, its response time to commands becomes more sluggish. The quality of service thus becomes marginal and user discontent is engendered; therefore, one can set forth a case for a computer which is dedicated exclusively to library applications. It appears that when one takes into account the storage and processing requirements for the information-retrieval and business administration sides of a computerized library, and the size of the associated library-user community, machines which are uniquely configured for library operations and reserved solely for library use may be in order.

The conventional devices for engaging on-line interactive computers are typewriter and cathode ray tube (CRT) consoles. In both, commands to the computer are placed by the operator through a typewriter-like keyboard; output information is obtained as a typewritten (hard) copy in the former device and as a television-like (transient or soft copy) display in the latter. Cathode ray tube displays write at a speed which is at least ten times the 100 words per minute rate of typewriter consoles and are therefore more desirable for man-machine dialog. It should be noted that a writing speed of 100 words per minute is below normal reading speed and hence imposes a drag on the reader. Hard copy output is inherent in typewriter consoles, whereas extra facilities in the form of either a high speed off-line printer or a separate dedicated display for hard copy purposes only must be provided when CRT consoles are employed. The need to have at least part of the output information ultimately available in hard copy form is universally accepted. The permanency and transportability of a written record make it mandatory to disseminate information in this form.

Although the superior writing rate of CRT displays gives them a clear advantage in an on-line interactive mode over typewriter consoles, several of their other attributes are controversial. Factors such as flicker, display brightness, crispness of characters, and character size

necessitate a variety of technical trade-offs in the design of these displays, with the result that they are not usually optimized with respect to all these factors. In addition, many designers resort to CRT's with a green phosphor as the electroluminescent screen material because of its high brightness. It is thought by some that the library-user community will find the greenish characters to be sufficiently annoying because of their strangeness to render CRT displays unacceptable. The alternative, a black-white, TV-type presentation is regarded by others as being too harsh and taxing on the eyesight to make it acceptable under conditions of sustained close-up usage.

Still others are skeptical of both classes of consoles because requesting information through a typewriter is considered to be unnatural. Time will take care of this matter handily, however. Typing skill among the young generation is widespread, and opportunities to use computers and typewriter consoles are increasing, not only in high schools, but in many elementary schools.

To be fully effective in library applications, consoles should incorporate certain design features which may not be necessary in other situations. It is generally agreed that the number of characters which a library console is capable of generating should exceed the standard ninety-six characters specified by the American Standard Code for Information Interchange (ASCII) set but it is difficult to get agreement on the exact number. The ASCII set, which includes upper- and lower-case Roman alphabet, Arabic numerals, and selected punctuation marks and symbols, should be supplemented by the Greek and Cyrillic alphabets (certainly the former), and additional symbols which are commonly employed in the literature of the mathematical and life sciences. When all reasonable possibilities are included, the number ranges from a minimum of two hundred to seven or eight hundred. Obviously the latter figure adds substantially to console costs.

Related to the size of the character-set is the size of the console keyboard. In addition to keys for ordering the character-set, one would like to include certain special purpose keys which can be struck in order to implement certain commonly used commands. Keys that initiate operations such as TRANSMIT FULL TEXT, TYPE AS SUPERSCRIPT (or SUBSCRIPT), and DELETE, avoid the need to spell out the command and thus simplify procedures. However, keyboard design involves a trade-off. Too few keys add to typing time; too many pose problems akin to those of a grand organ in that the keyboard becomes more difficult to master.

Still another consideration in console design for the library is the human factor. If library consoles are to be used by a substantial segment of a community, they must be adaptable to a variety of personal requirements. During sustained use of the console, the operator must feel comfortable and at ease at his station. The viewing screen should be adjustable to his eyesight characteristics, table space should be provided for writing and note taking, and the surrounding environment, including ambient lighting, should be conducive to easy reading of the material being displayed.

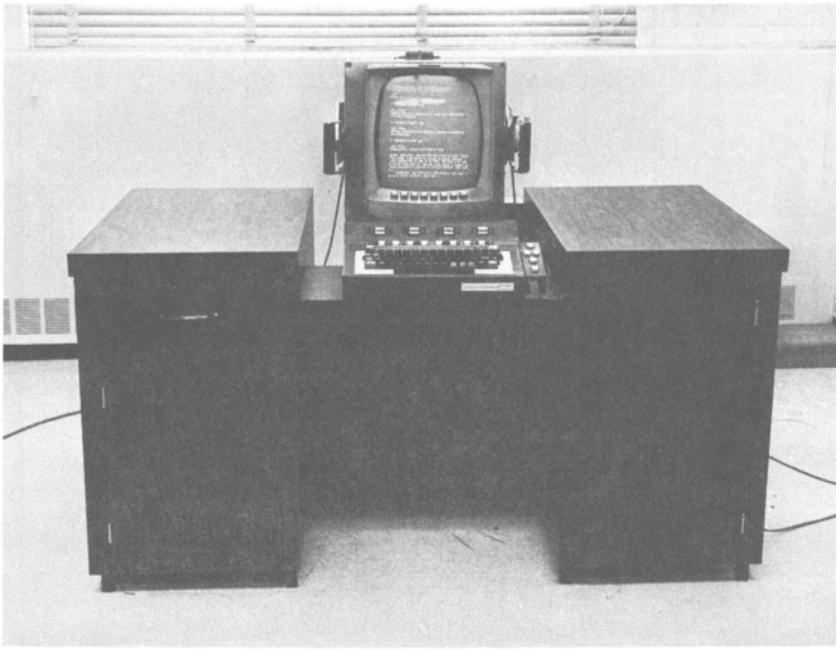

Figure 3. A Refreshed CRT Display Terminal

Figure 3 illustrates an experimental console being developed exclusively for library purposes by Project Intrex.* Its design has been made highly flexible so that user reactions can be observed as its special features are altered. The salient characteristics of the console are: a set of 192 characters; an ability to display a maximum of 1,736

* Intrex, standing for *In*formation *Tr*ansfer *Ex*periments, is a research project in library information transfer currently being conducted at the Massachusetts Institute of Technology, Cambridge, Massachusetts.

characters at a time; an ability to write superscripts and subscripts; a refreshment rate of the cathode ray tube display of fifty-seven frames per second in order to avoid annoying flicker; a set of status lights on the keyboard to advise the operator on the current progress of his requests; a set of wired function switches to enable the operator to move quickly back and forth among several display frames of information; a set of programmable logical function switches beneath the CRT display to apprise the operator of the choice of procedures currently open to him (the labels for these switches appear on the CRT directly above the switches); a sixty-key keyboard; and an operator's desk which includes writing and storage space, and facilities for raising, lowering, and tilting the viewing screen. It is expected that much insight into library console requirements will be gained from the reactions of users to this experimental console.

Substantial progress is being made toward storage of full text in microfilm form and an ability to retrieve the full text thus stored through commands executed from computer consoles. In the foreseeable future it appears that the full text of documents will have to be stored outside the computer rather than in storage devices associated with the computer. The combined information content of alphanumeric (letters and numbers) and pictorial text is too great to be handled by the types of computer storage devices that will probably be coming into service during the next five or so years. During this interval we must, therefore, look to non-computer techniques to solve current full-text storage problems.

It is generally agreed that there are currently problems in full-text access within the conventional library. It is mandatory that we find ways to reduce the sheer bulk of material that libraries must handle. In other words, information must be compressed into a smaller volume. Another problem is related to information retrieval itself. As the volume of materials that must be scanned or studied for relevancy increases, rapid accessibility to these materials becomes increasingly important. Furthermore, in a machine-oriented library, full text should be readily available at or near the consoles used to interrogate computer-stored catalogs. Thus, the requirement for high-density storage imposes a need for a medium which provides more compactness than does the printed page; the rapid accessibility feature implies a non-circulating store of full text; and the console-availability feature dictates either electrical transmission of text or duplication of full text storage facilities at each console.

In their attempts to meet these requirements, technologists turn to storage media which involve either magnetics or some form of optics, and in both of these, they may store information in either image or digitally encoded forms. Image storage means that letters, numbers, drawings and pictures are preserved as variations in the density of the storage medium, whereas digital storage implies that all information is stored as patterns of coded signals. In digital storage the patterns are usually constructed in accordance with the binary number system—that is, as patterns of signals having amplitudes of either *zero* or *one unit.* No matter which form of storage is used, the storage medium must meet rigid requirements of long life, minimum deterioration with age, durability in the face of repeated handling, and reasonable cost.

Magnetic storage may be accomplished through use of tapes, disks, wires, drums or cards upon which a magnetic material has been bonded, and each of these forms may be used to store information independently of an associated computer. Optical storage involves either traditional photographic storage or holography. The latter technique, although still in the research stage, offers excellent promise as a mass-storage device if it can be reduced to practice.

It is insufficient to evaluate each of these storage media by themselves; they must be considered together with the equipment used for displaying the stored information to the user. Ideally, display equipment should reconstruct the text, including drawings and photographs, with the same degree of resolution and contrast as provided by a printed paper copy. At present, only photographic processes can meet the combined requirements of high resolution and contrast, low cost, rapid reproduction and permanent hard copy, and these requirements are easily met only through the use of wet photographic processes. Electronic video recording and playback offers an interesting possibility for information storage and retrieval but, unfortunately this technique does not meet the total requirements set forth above for the library application. In addition to short life and questionable durability, magnetic tape recording equipment presently fails to meet the resolution requirement for high quality reproduction of pictorial images. Experiments have shown that text which is reconstructed on a line-by-line basis (that is, by line-scanning the original) should contain at least 1,000 line pairs per page, minimum, and preferably 1,500 line pairs per page. A line pair consists of a single full-density line followed by a single minimum-density line. The 1,500 line-pair require-

ment is mandatory when the original quality of text is poor or when very small type-size has been used for the subscripts and superscripts frequently found in texts dealing with mathematics and physical sciences. These figures correspond to a resolution requirement of 100 and 150 line pairs per inch for a ten-inch page, or approximately four and six line pairs per millimeter. For comparison, the upper limit of resolution of photographic film is in excess of 400 line pairs per millimeter.

Microphotography as a means of reducing storage space for library materials is thoroughly developed and has been for many years. However, its general acceptance has been slow to come, apparently for three reasons: lack of standardization on the reduction factor and physical form of the microfilm, dissatisfaction of users with microfilm viewers, and inadequate microfilm-to-hard copy printers. Very little progress is evident on standardization but some progress is appearing in the realm of microfilm readers and printers. Nevertheless there remains serious doubt whether the readership as a whole will ever subscribe with enthusiasm to the inherent constraints imposed on it by microfilm viewers. The superior reading comfort, flexibility and transportability of the full-size printed page places it in a preferred position. On the other hand, in the face of a cost of five cents to ten cents per page for hard copy, readers are likely to resort to a "first look" at film for decision-making, and then selectively move to hard copy for material of their choice. The lightweight, transportable, attaché case-size microfilm viewer that provides high quality images which can be comfortably viewed under sustained reading conditions without inducing eye strain, remains to be developed.

The central issue on microfilm-to-hard copy printers centers on the degree to which the gray-level content of photographic illustrations must be produced. Obviously, the value of fine grain gray-level information in pictorial material depends on the importance of the role pictures play in each document. Where two-level black-white reproduction is adequate, microfilm-to-hard copy printers which employ electrostatic printing techniques are now available on the open market. Full gray-level rendering, however, still requires the use of optical photography. Rapid strides are being made in the development of dry-process photographic papers. Here microfilm-to-hard copy printing will benefit from the emphasis currently being placed on dry papers that can be sensitized from cathode ray tube displays, particularly those that have a low-level light output. Dry-process papers have

traditionally been insensitive (their ASA numbers have ranged from 10^{-3} to 10^{-2}), but at least one dry silver paper with an ASA rating of approximately 1 is in the final product-development stage.

As library users turn to on-line computers as look-up aids, they inevitably will expect rapid accessibility to the full text of documents. Guaranteeing delivery of full text in soft copy, film, or hard copy page form within a few tens of seconds at, or near, user consoles appears to be a valid goal for future library service, and the technology for doing this now exists. Such a system might operate as follows: a user seated at his computer console desires to see the full text of a document so he calls for the document by typing its identifying number. As a result, the document, which is stored on some form of microfilm, let us say microfiche, is automatically located in its storage device, the first page is automatically moved into position and electronically scanned. The electrical image of the page is given the user's station code and then transmitted over a broadband transmission line to a viewing tube at the user's location. The viewing tube captures the page image and holds it until the user calls for a new page. Should the user decide he wants a permanent rather than a transient (soft) copy, he may ask for a film strip of his document; a hard-copy print can be made from the film strip if he so desires. Typical times for these operations which have been demonstrated experimentally are: time to display first page of text on cathode ray tube, nine seconds, maximum; to display each succeeding page, five seconds each; to obtain a 35 m.m. film strip of a five-page journal article, ninety seconds; and hard copy printing from film, seven seconds per page.

Although the above set of procedures has been demonstrated experimentally, practical design factors are presently preventing quick implementation of full-scale operational versions of such a system. For economical operation, the storage, retrieval, and electronic-scan equipment should be time-shared by all user stations. Automatic storage and retrieval devices which operate under computer control, provide the electronic scanning and transmission to a multiplicity of user stations, and accommodate any one of the several common types of microforms—roll film, fiche, cards, strips—are technologically feasible but unavailable as off-the-shelf items. Equipment of this kind will be custom designed to individual specifications until uniform standards can be agreed upon and established.

In order to minimize out-of-storage time of full text and to avoid long user queues, each microfilm frame should be scanned only once.

Repetitive scanning, in the manner of conventional television, ties up material too long to permit achievement of guaranteed accessibility. For one-shot scanning, therefore, we are forced to a storage medium at each user terminal. Cathode ray storage tubes show excellent promise for this purpose but they need further improvements for the library application. Only three-quarter size pages (approximately 6 by 9 inches) can now be displayed, and resolution is, at best, 400 line pairs per page. This resolution is less by a factor of at least three or four than minimum acceptable requirements for crisp, clear portrayal of text.

In contrast, equipment for film terminals is in good shape. Sixty-second film development time is easy to achieve, and it is within the state of the art to cut this time by a factor of two or three if necessary. Since film terminals are substantially more expensive than storage tube display terminals, it is likely that a single film terminal will serve several user stations.

Transmission of full text over wideband transmission lines imposes more economic than technological problems. Locally, textual images can be transmitted over distances of one-half mile or so simply by wiring the environment on a point-to-point basis with co-axial cable. Beyond that distance re-amplification of the signals is necessary to insure favorable signal-to-noise ratios. Signal routing through common-carrier telephone facilities is unnecessary if transmissions are confined within a private community, but for communication across public right of ways, either telephone facilities or privately-owned micro-wave links may be employed. Phone company facilities for long-distance interurban transmission of wideband signals are operational or under active development. Bell Telephone's T-1 carrier system for intercity communication under fifty miles digitizes all information prior to transmission and accommodates roughly 1.5 million bits of information per second. The T-1 carrier system is operational. An advanced system, T-2, for distances under 500 miles, will increase the transmission rate at least threefold and should be operational in the early 1970's.

Optical character readers (OCR's) show promise as devices for digitally encoding printed alphanumeric materials. Their merit lies in their potential ability to encode faster and more accurately than keypunch operators. Their reading speed depends upon the particular piece of equipment, but may be as high as 2,000 characters per second. Multifont-reading capability is also possible, but the handling

of subscripts and superscripts poses a formidable problem. Because of the cost factor, it may be several years before OCR's can compete economically with the manual-keypunching process.

Interest in OCR's stems from their possible use for converting printed matter into digital form. Equally important is the digital encoding of new material as it is generated. If one accepts the premise that printed documents will continue to exist in the foreseeable future, then a valid goal is to obtain a digitally encoded version of a document as a natural fallout at some point in the original document fabrication process. For example, the author might be expected to provide descriptive cataloging and subject indexing information upon submission of his manuscript for publication, or to furnish a digitally encoded version of his typewritten manuscript for automatic cataloging and indexing purposes. Alternatively, professional societies and book publishers might make available through their printers certain digitally encoded information as a natural by-product of a computer-controlled typesetting process. Beyond the fact that an inexpensive digital attachment to the common typewriter does not exist, it should be evident that important economic and social elements are involved in the direct procurement of digitized copy coincidently with document manufacture. These issues will, however, have to be resolved if large quantities of information are to flow inexpensively into the digital domain.

Since, in the computer domain, information exists in electrical-signal form, it is naturally suited for transmission over wires and through space. In principle, therefore, it is unnecessary to duplicate digitized information at every library; each can specialize in a different field of information and each can draw on one another's data bank through communication channels as requests occur.

Figure 4 shows a network of three computers, at each of which five consoles are available. A user at each console may gain access to each computer through his local telephone switchboard. Connection between a switchboard and a geographically remote computer may be either on a dial-up basis or over a leased line. If several terminals in a cluster must share a single leased line, a possible queueing problem can occur. Since each console terminal connects to three machines, the data format, the retrieval programs, and the manner of addressing the machines must be either identical in order that only one set of procedures must be mastered, or the user population must learn separate procedures for engaging different machines. Thus, if

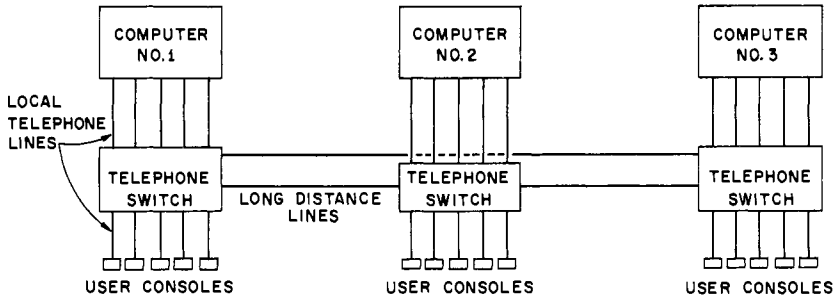

Figure 4. An Information-Transfer Network in which Each User Console is Connected Directly to Each Computer through a Telephone Switch

the network configuration of Figure 4 is employed, standardization among information retrieval systems is of prime importance from the viewpoint of ease of operation.

Figure 5 illustrates an information-transfer network configuration in which a local buffer computer is interposed between user and central computing facility. The buffer machine may serve to route requests to the proper central machine and to condition, or reformat, requests so that they will be recognizable in a heterogeneous machine environment. Thus, in Figure 5, the buffer machines serve as message handlers or "information brokers" between the users and the main machines.

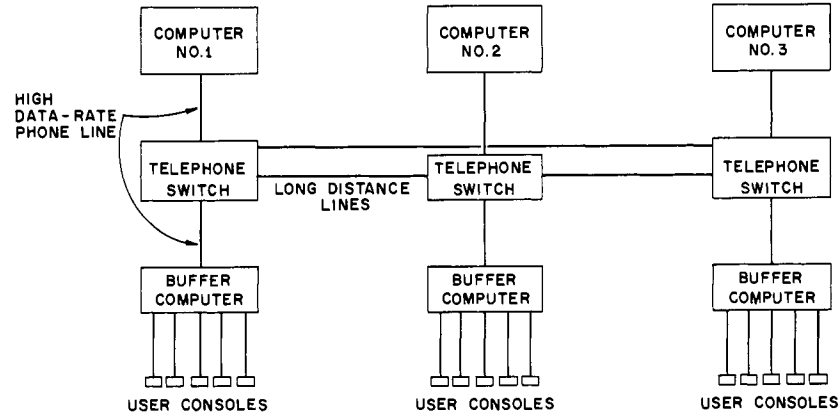

Figure 5. An Information-Transfer Network in which Each Console is Connected to the Main Computer through a Buffer Computer and a Telephone Switch

When a request for information is given to a buffer machine, it directs the request to the central computer which contains the information. In the event the main machines are not identical and the information in each is formatted differently, the buffers also serve to reformat requests. These capabilities come, of course, at the expense of complexity at the buffer computer level. Note also in Figure 5 that single high data-rate lines are suggested rather than multiple low data-rate lines.

A modification of the network shown in Figure 5 would permit local information retrieval to be handled directly by the local central machine. Only when requests must be filled by reaching out to another data base stored in a geographically remote machine would these requests be referred to the buffer machine. Thus the load on each buffer would be reduced by the amount of local traffic at each location.

With respect to the telephone, there is no technical obstacle to the interconnecting of computers in any of the configurations illustrated. However, since leased lines are expensive and the alternative, on-demand service, is subject to prolonged delays during peak telephone-load periods, it appears that approaches which minimize long distance communication will ultimately be preferred.

As yet, no exhaustive studies are available which resolve trade-offs among the various technological and economic factors involved in computer-library networks. Much valuable information should be gained shortly, however, from a rather ambitious program now being sponsored by the Advanced Research Projects Agency (ARPA) of the United States Government. ARPA is in the process of establishing a coast-to-coast computer communications network among several of its major contractors. The network, which should be at least partly operable by 1970, will link together an assortment of "multi-lingual" machines on a time-sharing basis. Through this network valuable insights should be gained on matters pertaining to load sharing, information transmission, and data sharing among machines that are geographically remote from one another.

Information networks raise issues beyond those that are purely technological. The very existence of a library network necessitates a commitment to the needs and requirements of the total population which the network serves. Such a commitment may be contrary to current policies of the autonomous library. For example, the matter of priority in a queue of users seeking service is a delicate one. Do

local users have first call on locally stored information, or are all users in the net, local and remote, to be handled equally on a first-come first-served basis? Agreement must also be reached on the relative sizes of the data bases and computing facilities at the various nodes of the network, and these matters bear on others such as computer load sharing, computer and communications cost sharing, data base up-dating, and computer program maintenance and upgrading.

In summary, a substantial body of computer, communication and microphotographic technologies are available for aiding business and informational-retrieval processes in libraries. Nevertheless much research and development work on new technologies must continue if the major functions of library systems are to be executed principally through the use of electronic equipments.

Progress Toward a National Serials Data System

SCOTT ADAMS

WHERE DO THE responsibilities for the planning, development, and management of a national serials data system lie? With the federal government? If so, through what mechanisms? Will such a national system come about through the voluntary merger of existing local, state and regional systems agreeing to submit themselves to standards and controls in their own interests, or will it require the imposition of an independent authority?

As individual librarians have applied computer technology to serials management within their libraries, and as these libraries have become associated in local, state and regional groupings, major capital investments are being made, and independent management, policy and fiscal entities are being established. Are we on the way to creating a national serials "unsystem" rather than a national system, and when shall we pass the point of no return?

The purpose of this article is to review recent trends in one particular area of computer application and systems development important to libraries, that of serials data, in an effort to find answers to some of these questions. At the outset, I should like to provide a working definition of the words "system" and "network" as I shall use them. A "system" may be defined arbitrarily, as a confederation of interrelated functions under a single control leading to the production of useful products or services. A "system" may arbitrarily be distinguished from a "network" in that the latter lacks a single control function.[1] A "network" consists of independent or quasi-independent functional units, with interrelated functions and services and cooperatively-agreed upon standards, acting in a decentralized but reticulated mode. By this definition, one may refer to a municipal or county library system, while

Scott Adams is Special Assistant to the Foreign Secretary, National Academy of Sciences, Washington, D.C.

an interlinkage regionally of functioning units under multiple jurisdictions, such as exists for purposes of interlibrary loan, constitutes a network. As the word is used in this paper, a serials data system would imply multiple units under some form of common control and producing a common product or service. A serials data network, on the other hand, would imply loosely associated, cooperative functions and services under independent controls.

It may be arguable whether the functions and services implied by this paper require a national system, or whether they can be accomplished through networking. My *a priori* position is that in order economically to continue the serials location function performed historically by the published *Union List of Serials*, a system under some form of centralized management is required. The discussion to follow will demonstrate this bias.

The Union List of Serials

While systems in other areas (telephone, railroads) have demonstrated historically a pattern of merger and growth, the evolution of a national serials data system has demonstrated a reverse situation. A system assembled by librarians for producing a product (the *Union List of Serials*) and for providing a national service (location for interlibrary loans) has broken down into a multiplicity of local efforts of questionable compatibility. While some evidence exists that larger systems are evolving, doubt remains that a national system can be reconstituted.

The initial creation by the library community of the first edition (1927) of the *Union List of Serials* represented a national and even international integration of local efforts to create a serials location system. A list of serial titles held by libraries of the Chicago area (1901) grew through the participation of the John Crerar Library to a list of 12,000 titles in 1906. In 1916, this Chicago list became the basis of a North Central Union List which consolidated the holdings of principal research libraries in the Central States.

After World War I, H. M. Lydenberg, working with an American Library Association committee, expanded the project still further to include all major research libraries in the United States. In 1925, that American Library Association committee agreed to include serial holdings of Canadian libraries. Under ALA sponsorship, the work of producing the first edition of the *Union List of Serials* under Winifred Gregory, as editor, and with H. W. Wilson Company as publisher,

took four years. Funding, estimated at $36,000, was by advance subscription from participating libraries.

The history of the first edition of the *Union List of Serials* demonstrated merger and integration of effort. Standards governing bibliographic elements were established by committee, and the participants cooperated voluntarily to produce a product which would function to provide libraries and their users with two basic services: a) authoritative identification of serial titles and their sponsoring organizations, and b) location information to facilitate interlibrary loan.

The first edition produced a $10,000 surplus by mid-1930; this was used to fund two supplements covering serials to 1933. In 1936, the H. W. Wilson Company proposed a second edition. The American Library Association created a Committee on the Union List; $48,000 was advanced by the Rockefeller Foundation.

The administrative history of the second edition was less smooth than in the case of the first edition, but in 1943 a second edition of the *Union List of Serials* was published. This contained 100,000 titles in 365 pages and recorded holdings from 650 libraries.

Following its publication, a long series of discussions, investigations, and studies were initiated involving the processing department of the Library of Congress and the Joint Committee on the *Union List of Serials*. A central theme was the possibility of automating the serial recording function through the use of punched cards. To the late Alton H. Keller belongs the credit of first conceiving, in 1948, the application of automatic data processing to serial records. Keller, in planning the future of the Library of Congress's serial record, proposed to produce a flexoline file by punched cards, the file to contain the following bibliographic elements:

a) distinctive title entry,
b) place of publication,
c) beginning and closing date of publication,
d) frequency of issue,
e) Library of Congress numbers for bound issues,
f) division of Library of Congress responsible for custody and service,
g) Library of Congress holdings, and
h) indication of current receipt.

Upon completion of the Library of Congress' flexoline file, some 250,000 serial titles would be recorded. Keller suggested that "by using

these same methods, it should be possible to develop a Union Catalog of Serials on punched cards which would show the serial holdings of the larger American research libraries. This could be kept updated on a continuous basis, and from it, listings of publications and their locations could be prepared by subject, languages, country of publication, date, etc." [2] Subject control would be accomplished through assignment of the Dewey Decimal classification. Keller's proposal, well in advance of its day, even suggests that input from participating libraries be in machine-readable form.

The staff of the Library of Congress reviewed the Keller proposal over the next two years, identifying problem areas such as costs of editing the file, costs to participating libraries, and of course, the time involved in processing a file of an estimated 8 million cards. The Library of Congress staff presented this concept to a meeting of the Joint Committee on the *Union List of Serials* in June, 1952.

The forecasting of cost data benefited from the operating experience of *Serial Titles Newly Received* issued by the Library of Congress (1951-1952) which applied the flexoline-punched card technique proposed by Keller. An initial estimate of $935,539, involving editing, keypunching, publication, and the establishment of subject and country files, was increased to $1,095,682 by January, 1953.

The Keller proposal involved keypunching title and holding information from the *Union List of Serials* (supplemented by other sources), printout for offset publication of a union catalog of periodicals, and the maintenance of auxiliary files by subject (Dewey Decimal) and by country. The resulting files were considered exclusively as the base of a publication system; the objective was to produce a published *Union Catalog* and specialized listings; and the size of the card file precluded any consideration of search.

In the 1950's the Joint Committee on the *Union List of Serials* had three major problems before it: a) What to do about a third edition of the *Union List of Serials*? b) What to do about *Serial Titles Newly Received*? and c) What to do about the Library of Congress's proposal for a permanent *Union Catalog of Periodicals* on punched cards?

It resolved the first by seeking a grant from the Council of Library Resources to produce the third edition, subsequently published in 1966. It resolved the second by persuading the Library of Congress to broaden the base of its publication of *Serial Titles Newly Received* to include titles and holdings of other libraries. In 1953, this resulted in *New Serial Titles*. It resolved the third by seeking a grant from the

Rockefeller Foundation to support a study published in 1957 under the title, *A Permanent Program for the Union List of Serials.*[3] This report, recommended that a *Union Catalog of Serials,* based on punched cards, should be established in the Library of Congress to provide a basis for publication on national and regional union lists as well as special lists by subject fields and by country of origin. The estimated cost of this project was $2,673,222 of which $975,000 would be used to subsidize reporting by participating libraries. Funding of this magnitude was considered impracticable, and the proposal died. Until 1961, however, the Library of Congress continued to use punched cards in preparing *New Serial Titles.*

What can be learned from the foregoing effort to build a national union list of serials and to maintain it through the application of mechanization? First, it appears that all participants were preoccupied with the objective of a published union list. Little consideration was given to other means of providing the location functions which the publication served, despite the fact that as early as 1946 a proposal had been made that electric accounting machinery could be adapted to perform search functions.[4]

Second, this preoccupation derived in part from the level of data processing technology then available. During the period when the possibility of mechanization was discussed, punch cards represented the technical limits of library applications.

Third, costs and benefits were measured entirely from the point of view of preparing published lists. In the absence of estimates of the functional benefits a system might produce, were it capable otherwise of supporting the location search function, the benefits were underestimated.

Fourth, whereas the Joint Committee on the Union List of Serials and the Library of Congress cooperated successfully in the preparation of the third edition of the *Union List* and in the expansion of *New Serial Titles,* no clear resolution emerged as to the ultimate responsibility of continuing the national effort. The one exception to this was the acceptance of responsibility by the Library of Congress to produce *New Serial Titles* as a mechanism to update continuously the third edition of the *Union List.* It should be noted that this updating was a compromise in that it reported holdings for new titles by a limited number of libraries and did not reflect the changing national picture of retrospective and current holdings of serials across the country.

The Trend Toward Decentralization

Parallel with the effort to produce the *Union List of Serials* as a national location system, ran a tradition dating to the early days of this century of compiling local, regional, and special subject union lists. As early as 1901, the Library of Congress itself, published a union list of periodicals received by sixteen libraries in the District of Columbia.[5]

The appearance of three editions of the *Union List*, each increasing in scope of titles and of library holdings reported, had little effect on the proliferation of such local lists. The second edition of the *Union List* contains, as an appendix, an impressive bibliography of published local and regional union lists as of 1954.

Between 1940 and 1957, twenty-five union list projects were published, four-fifths of them regional or local.[6] Seventeen more were announced or completed by 1962.[7] Freitag, in 1964, updated the bibliography published in the second edition of the *Union Lists of Serials* [8] with a listing including 364 union lists published in the U.S. She reports a hundred new lists known to her office since the publication of the supplement.

The motivations for the production of such local and regional union lists appear to be various. In most instances they reflect a pattern of local library cooperation in the interlibrary loan process. Smaller libraries, excluded from reporting in the national *Union List*, have apparently been strongly motivated to publicize their willingness to share their holdings with their colleagues. In the special library field, for instance, a different category of libraries, industrial and commercial, has been concerned with developing cooperation. Finally, the device of a union list has had an appeal to library systems operating under a single jurisdiction such as a university library complex, both to facilitate reader access to the total library resources of the university and to avoid unintentional duplication in acquisitions. A large number of these local and regional union lists have resulted from the voluntary professional activities of library associations in pursuit of cooperation as a professional goal. Whatever the motivation, the publication of local and regional union lists has been a part of the American library tradition of interlibrary cooperation.

The application of data processing technology to bibliographic information about serials has built on the foundation of these uncoordinated local efforts rather than on the national tradition. It is historic irony that proliferation of local computer applications trod hard on

the heels of the decision to reject data processing technology for the production of a national union list.

Without tracing in detail the history of data processing applications to serials control, it may be noted that the National Reactor Testing Station Technical Library at Idaho Falls, had a union list of periodical holdings of six libraries by 1960. The influence of the Library of the Advanced Systems Development Division, IBM and the IBM-Endicott Library in pioneering applications for serials control, may be noted. As early as 1961, McCann reported on five special libraries which had developed data processing systems for purposes of serial subscription renewals, listing, routing and claiming.[9]

A prototype application which attracted wide attention was at the University of California, San Diego, where a pilot operation was completed in 1962, and a system covering 5,000 titles became operational in 1964.[10] The development of an "anticipated arrival" card as the basis of check-in and claiming, was widely copied in other applications. A primary purpose of the San Diego project was to produce a union list for all campus libraries at the University.

Subsequent applications proliferated in the early 1960's with U. of Illinois (Chicago),[11] the National Science Library of Canada,[12] the School of Medicine at Washington University,[13] Purdue,[14] and many others developing systems independently. As of September 1, 1969, the Information Systems Office, Library of Congress, had a record of 300 American institutions or groups of institutions utilizing data processing equipment for the purpose of serials control. In a large number of these cases, the application was directed toward the production of a published list of serial titles currently received or held by one or more units of a library system.

Purposes common to these pioneer efforts appear to be: 1) to improve the processing of serials (ordering, recording, claiming, binding), and 2) to produce published listings of holdings and locations of one or more library units for staff and patron use. The proliferation of individual systems efforts during these pioneer days may be attributed to two circumstances which may be inferred from published accounts. First, the complexity of serials and their processing have long provided the profession with an intellectual challenge; experimentation with computer technology to solve long-standing problems offered exciting possibilities to the adventurous. Second, this type of application represented to the profession a prime opportunity for learning through doing. The significance of this educational experience should not be discounted.

On the other hand, the proliferation of these independently planned systems has been costly; repetitious conversion and programming could have been avoided if a common data base had been in existence. Further, the emphasis was usually placed on the development of a system to support locally the serials processing function, not the location function.

The decision to include data elements in these serial lists appears to have been arrived at independently. Some conformity has, however, been achieved, first through the influence of prototype systems such as the one at the University of California at San Diego, and second, through financial limitations which reduced the number of elements to a minimum.

The Trend Toward Regionalization

The production of individual lists by individual institutions represents the initial application of data processing to serials control. Over the past few years there has emerged a second level effort to develop statewide or regional serials data systems. While the majority of these state or regional projects are in a planning or a developmental stage, and while a comprehensive listing would be difficult to compile, notes on representative proposed developments may be offered to indicate the trend. Excepted from these examples and deferred for later discussion are instances of serials data systems of national pretension.

Illinois. The Illinois State Library has under development a statewide union list of serials planned to incorporate the holdings reported by approximately thirty small college libraries in the state. A 1965 union list of serials of the University of Illinois Libraries has been used as the basis of the list.

Indiana. An initial project undertaken by the four state university libraries (Indiana University, Ball State University, Indiana State University and Purdue), with funding from Title III of the Library Services and Construction Act, has been expanded to include the holdings of sixty-two public, academic and special libraries of Indiana and the Indiana State Library. This project is conceived of as a statewide serials data bank, capable of association with any national network to be developed.

New York. Two projects of statewide scope are under way. Led by the SUNY Syracuse Biomedical Communications Network Group, sixty library components of the State University of New York (SUNY)

have cooperated to publish a union list recording holdings for 25,000 serial titles.[15] This list includes all subjects except law. The New York State Library has developed plans for a New York State Union List of Serials under the three R's program. This statewide list would build on the base of the SUNY list and would absorb other regional union lists which have been developed in New York State. Its scope in general would be the same as that of the third edition of *Union List of Serials;* it would, in a phased program, cover all the major resource libraries of the state, public and private, and would be designed to support the New York State Interlibrary Loan Program (NYSILL). Plans call for search as well as for publication capabilities.

Ohio. In 1968, Wright State University published a union list of 8,880 titles held by ten university and college libraries and twenty-nine special and public libraries in the Miami Valley of Ohio.[16] In its research and development program designed to increase the availability of library resources for use in the educational programs of Ohio colleges and universities, the Ohio College Library Center has assigned priority to a mechanized shared-cataloging system. It has defined as a later objective, work on a serials control system to be designed to facilitate library control of serials holdings in the state of Ohio.

Oklahoma. A list of current periodicals held by the Oklahoma State University and the University of Oklahoma has served as the basis for a union list of serials, scheduled for publication in 1970, which represent the holdings of sixteen libraries.

Oregon. The nine institutions comprising the Oregon State System of Higher Education have a union list of serials in an early stage of development. The data base and programs for this effort were acquired from the Union Catalog of Medical Periodicals System.

Washington. Under the leadership of the Washington State Library, a three-step serials control system is being planned. The first step envisions a statewide listing of serial titles held by Washington libraries; the second, the development of a serials control system for the Washington State Library and its branches; and the third, the production of a union list of serials with specific holding information for all libraries in the Washington State Library network.

This is by no means a comprehensive list of state or regional projects in the planning or early developmental stage. As representative samples, however, they do indicate a trend toward the development

of larger cooperating units, away from the individual library project.

Several generalized comments, based on personal correspondence with the planners of these statewide or regional systems, may be made:

1. This trend has been abetted by funding from Title III of the Library Services and Construction Act amendments of 1966, which support the cooperative development of interlibrary loan networks.
2. Nearly all these systems express need and desire to be compatible with any national standards and any national system to be established.
3. In the absence to date of national standards (e.g., minimal data elements to be used for purposes of identification and location), reactions vary from deliberate deferral of projects, through cautious step-by-step activity, to a do-it-yourself philosophy. As one librarian states:

> Needless to say, a national serials system would ideally originate from the top and a single data format would be utilized by all the libraries desiring to go into it. . . . Libraries have traditionally developed their own systems without regard to compatibility and interchange for so long that it might be difficult to alter this pattern. Much of the reason for this pattern has been that higher level guidance has been too slow and individual libraries and librarians have been too progressive to wait indefinitely for such assistance.

4. No librarian responsible for the development of these regional systems has proposed a master plan for a national serials system, although several have expressed a hope that their own systems might be considered a prototype for such a plan.

The National Level

Present efforts to achieve a national serials data system owe their impetus not to the library community, but to considerations at the level of governmental science information policy. As it became acquainted with the problem of document location, COSATI identified the need for a national system which would locate serial publications in science and technology.

The System Development Corporation's *National Document-Handling Systems in Science and Technology* proposed that the federal government assume responsibility for assuring the existence in the U.S. of at least one accessible copy of each significant publication in

science and technology. It noted that this responsibility extended to the holdings of private and university libraries, that a "national union listing and an indexing of document holdings of major libraries" was required, and further that "among the most serious needs that are not now being fulfilled within the library community today is that of maintaining a union list of serials." [17]

COSATI discussions about the need for a union list system to aid in locating serials in science and technology led to a National Science Foundation (NSF) contract with Information Dynamics Corporation (IDC) to study the feasibility of a union list in machine-readable form.[18]

The IDC study, while concentrating more on the feasibility of alternate routes to building a large national data base than on the organization and performance characteristics of a location system, concluded that the mechanization of a union list of serials in science and technology was economically feasible.

Despite the strong program interests of the National Science Foundation and the Council on Library Resources, both of whom were well-disposed to funding further work toward a national serials data system, a two-year period of inaction followed. Ultimately, the Joint Committee on the Union List of Serials established a subcommittee to draft a developmental proposal, and Dr. F. H. Wagman, its chairman and concurrently chairman of an American Research Libraries Ad Hoc Committee, reported in January, 1967, at a meeting on the Joint Committee on the outlines of a three-phase proposal for a National Serials Data Program to be centered in the Library of Congress.

The National Library of Medicine and the National Agricultural Library expressed strong interest in contributing both to the support and to the conduct of the program. As a result, the proposal became absorbed in the more comprehensive plans of the three national libraries to coordinate the mechanization of their bibliographic processing functions. A public announcement of this cooperative effort was made in July, 1967, and a Task Force was appointed to implement specific projects, including the serials data program. Composite funding from the National Science Foundation, the Council on Library Resources, the National Library of Medicine, the National Agricultural Library and the Library of Congress was arranged for the first phase. To the Library of Congress was assigned responsibility as executive agent.

The Information Systems Office of the Library of Congress em-

ployed Thomas Nelson Associates to survey libraries to assist in a determination of the relative utility of 278 bibliographic data elements relating to serial publications. Based on sampling the frequency of consultation of these elements, estimates were made of the magnitude of the file maintenance requirements. The information was consolidated and reported to the Joint Committee and the directors of the three national libraries.

Noting that the selection of data elements represented multiple functions (identification and location, the processing of serials, and the study of the intrinsic characteristics of serials as a form), the three directors decided to limit the second phase of the project in two ways: by placing major emphasis on bibliographic elements required for identification and for holdings information, and by limiting the field to serial publications in science and technology.

Responsibility for the second phase, or pilot project was assigned to the National Agricultural Library. The National Agricultural Library has contracted with the Association of Research Libraries to provide a project leader and staff, and the pilot project started in July, 1969. The Task Force of the three national libraries continues to have an advisory role.

Much basic fact finding was accomplished during the first phase. The exploration of desirable data elements led to the development of a MARC format for serial publications,[19] and many problems were identified and resolved. Still lacking in this effort, however, is a simple, standardized format for the identification and location functions performed historically by the *Union List of Serials.* It is to be hoped that the pilot project will stabilize this format, thereby establishing a model to which the developing local and regional systems can conform.

Still lacking also is an adequate conceptualization of a search service system to provide nationally for the location function historically performed by the *Union List of Serials.*[20] Attention has consistently been focused on the development of systems to produce printed products, rather than on systems to facilitate the search of files to provide holdings information on specific titles. There is need to create models of various configurations of local, regional, and national serials data files in order to design an optimum system to provide a national search and location function.

The National Serials Data Program undertaken by the three national libraries, while representing a lineal descendant of the *Union*

List of Serials, is but one of three developmental serials systems inviting participation at the national level. The other two are *ACCESS,* developed by Chemical Abstracts Service, and the *Union Catalog of Medical Periodicals,* developed by the Medical Library Center of New York.

ACCESS

Since 1922, the most widely used list of serials in a broad subject field has been the quinquennial list of periodicals abstracted by *Chemical Abstracts.* In October 1969, the Chemical Abstracts Service published *ACCESS* from a computerized data base. *ACCESS* is a vastly expanded version of this established service which was initially designed to help chemists locate within libraries the full text of the papers covered by *Chemical Abstracts.*

The first edition of *ACCESS* contains over 16,000 entries for serials and 4,500 entries for monographs, in addition to a number of chemical journals which pre-existed *Chemical Abstracts.*[21] Holdings data, totalling some 727,000 locations, have been supplied by 325 libraries in the United States, and seventy-four libraries in twenty-eight other countries.

The bibliographic data elements average twenty-four for serial entries and twenty-five for the non-serials. Of interest to librarians is the fact that entry both by direct title in the original language of publication and by ALA cataloging rules are included in the elements for serials. The data base used to produce *ACCESS* will be used to produce quarterly supplements and subsequent editions of the published list and will also be searchable, so that lists of journals by language, country of publication, frequency of publication, type of journal, and other parameters can be produced.

ACCESS represents a significantly large data base in the scientific and technical serials universe. Comparison studies between *ACCESS* and the lists of serials covered by nine other indexing and abstracting services in science and technology have shown that *ACCESS* includes from 29 percent (agriculture) to 82 percent (nuclear sciences) of the serials processed by other services.

Union Catalog of Medical Periodicals

As in the case of *ACCESS,* the *Union Catalog of Medical Periodicals* (*UCMP*), developed by the Medical Library Center of New York, covers the serial literature of a broad subject area in science and tech-

nology. Unlike *ACCESS*, it is an open-ended system, capable of expansion and use at multiple locations, rather than a centralized operation under single management.[22]

With a data base of approximately 15,000 titles in the medical and paramedical sciences, a limited number of bibliographic elements (nine), a thoroughly tested package of computer programs, the *UCMP* owes its strength to its simplicity and wide availability. Any library or group of libraries may acquire the serials data base stripped of holdings information, post additional holdings data, and add titles, and create its own internal or regional union list of serials.

The national use of the *UCMP* format and programs has grown rapidly over the past few years. Starting in 1967 with a group of eighteen medical libraries on Long Island, the system has expanded both geographically and in its subject coverage. Nine principal medical libraries of Virginia, North and South Carolina and Kentucky (the VINSCKY group) were followed by twenty-two Texas health science libraries in producing *UCMP*-based regional union lists. Rutgers University in 1968-69 broadened the *UCMP* data base to develop a university-wide union list of serial holdings, and together with the New Jersey State Library has proposed the use of the format for a statewide list. Eight medical school libraries in Missouri, Kansas, Nebraska, Colorado and Utah have used it to produce a union list of medical serials for the Central States Regional Medical Library Group, and as mentioned earlier, the format has also been proposed for statewide use in Oregon.

The *UCMP* system has obvious appeal in that it is simple, easy to implement and operate. It is modest in cost, and it works—at least for the purpose of producing published lists.

Discussion

In these early years of computer application to serials, the following trends can be noted:

1. Concurrent with the abandonment of the Keller proposal to base a national union list of serials on a punched card file, individual library applications began to proliferate.
2. The local applications are now being superseded by statewide and regional systems.
3. No national system with the same universality as the *Union List of Serials* has yet emerged. Three systems of national potential are under development in areas of science and technology.

This situation raises a number of questions for the library profession and, even more importantly, for library users. The first question is: Do we want a national system for purposes of locating serials holdings? Some confusion of priorities exists between local systems to facilitate the processing of serials in individual libraries, and regional and national systems created to satisfy location requirements.

If the profession agrees on assigning priority to a location function, should this function be accomplished by publication as heretofore? If so, should the concept of a comprehensive national union list of serials be revived, or should we depend on multiple local and regional published union lists?

Is it feasible to design a location system which would depend on search of machine-readable files rather than on published products to accomplish the location function? How can we find out? Who would design, develop, and manage such a system? Could it be used to produce multiple published union lists, as well as to provide search services? What other benefits could such a system provide?

Finally, to return to the questions raised in the opening paragraph, where does the responsibility for leadership lie? in the library profession, through the Joint Committee on the Union List of Serials or the Association of Research Libraries? in the federal government, through the Library of Congress, the three national libraries task force, or through the Office of Education?

If a national serials data system for location purposes is to emerge in the foreseeable future, these are all questions which call for earnest consideration and prompt answers. Otherwise we shall be burdened for years to come with fragmentary and partially compatible bits and pieces of an uneconomical network, frustrating to use and entrenched in practice. So far our approach to this question has been the reverse of that recommended by a former Librarian of Congress, Luther H. Evans, that "the matter [of union lists] should be dealt with first on a national basis before remnants of the task are left for regional solutions." [23]

References

1. Weinstock, Melvin. "Network Concepts in Scientific and Technical Libraries," *Special Libraries*, 58:328-34, May 1967.

2. Keller, Alton H. "A Plan for the Development of a Union Catalogue of Serials by Punched Card Methods." Washington, D.C., Library of Congress, Jan. 16, 1950. (Unpublished.)

3. Joint Committee on the Union List of Serials. *A Permanent Program for the Union List of Serials.* Washington, D.C., Library of Congress, 1957.

4. Gull, C. D. "The Compilation and Production of Indexes to Catalog and Bibliographies by Punched Card Machines." Washington, D.C., Library of Congress, 1946. (Unpublished.)

5. Library of Congress. Division of Bibliography. *A Union List of Periodicals, Transactions, and Allied Publications Currently Received in the Principal Libraries of the District of Columbia.* Washington, D.C., U.S.G.P.O., 1901. Also of historical note is the *Cooperative List of Periodical Literature in Libraries of Central California.* 3d ed. Berkeley, 1902. The first edition, 1880, included the holdings of nine Bay area libraries; the second, 1892, twelve; and the third, eighteen.

6. Joint Committee on the Union List of Serials, *op cit.*, p. 9.

7. Galvin, Thomas J. "Regional Union Lists, Some Unanswered Questions," *Library Resources and Technical Services,* 8:5-14, Winter 1964.

8. Library of Congress. General Reference and Bibliography Division. *Union Lists of Serials; A Bibliography.* Compiled by Ruth S. Freitag. Washington, D.C., Library of Congress, 1964.

9. McCann, Anne. "Applications of Machines to Library Techniques: Periodicals," *American Documentation,* 12:260-65, Oct. 1961.

10. Vdovin, George, *et al.* "Computer Processing of Serial Records," *Library Resources and Technical Services,* 7:71-80, Winter 1963.

11. Schultheiss, Louis A., *et al. Advanced Data Processing in the University Library.* New York, Scarecrow Press, 1962.

12. Brown, Jack E. and Wolters, Peter. "Mechanized Listing of Serials at the National Research Council Library," *Canadian Library,* 19:420-26, May 1963.

13. Pizer, Irwin H., *et al.* "Mechanization of Library Procedures in the Medium-Sized Medical Library: The Serial Record," *Medical Library Association Bulletin,* 51:313-38, July 1963.

14. Hammer, Donald P. "Reflections on the Development of an Automated Serials System," *Library Resources and Technical Services,* 9:225-30, Spring 1965.

15. SUNY Biomedical Communications Network. *SUNY Union List of Serials.* 2d ed. Syracuse, State University of New York, 1967.

16. Shields, Joyce, ed. *Union List of Serials in the Libraries in the Miami Valley.* Dayton, Wright State University Library, 1968.

17. Carter, Launor F., *et al. National Document-Handling Systems in Science and Technology* (Information Sciences Series). New York, Wiley, 1967, pp. 37, 119.

18. Creager, William A., *et al. A Serials Data Program for Science and Technology; Results of a Feasibility Study.* Reading, Mass., Information Dynamics Corp., 1965.

19. U.S. Library of Congress. Information Systems Office. *Serials: A MARC Format.* Washington, D.C., Library of Congress, 1969.

20. For a preliminary conceptualization of a regionally decentralized system see: Adams, Scott. "Proposal for a National Serials Data System." Bethesda, Md., The author, 1969. (Unpublished.)

21. From information supplied by Chemical Abstracts Service, Columbus, Ohio.

22. Felter, Jacqueline W. "Initiating a Mechanized Union Catalog for Medical Libraries in Metropolitan New York," *Special Libraries*, 55:621-24, Nov. 1964; and Sprinkle, Michael D. "Regional Utilization of the Union Catalog of Medical Periodicals System," *Medical Library Association Bulletin*, 57:244-49, July 1969.

23. Evans, Luther H., *et al. In* Harold Orlans, ed., *Federal Departmental Libraries; A Summary Report of a Survey and a Conference*. Washington, D.C., Brookings Institution, 1963, p. 24.

A National Bibliographic Data Base in Machine-Readable Form: Progress and Prospects

RICHARD DE GENNARO

LIBRARIANS ARE gaining experience with localized computer systems, they are struggling with the problem of how to integrate the use of MARC data into their technical processing operations, and they are contemplating the intriguing possibilities of a national library network. As they do so, they are becoming increasingly aware of the necessity for converting their retrospective catalog records to machine-readable form which will be the foundation of the complex automated systems that the future requires. This article will address the question of why retrospective conversion is necessary, and it will attempt to show that it is a feasible objective by citing significant research and recent, continuing large-scale conversion projects. It will explore the means by which retrospective conversion might be accomplished as well as cost and time projections. Emphasis will be placed on the Library of Congress' current and forthcoming activities in this area because they are of particular significance in the creation of any national bibliographic data base in machine-readable form.

Why is retrospective conversion necessary? Most librarians have accepted the idea that conversion of current and future catalog records to machine-readable form is both a desirable and a necessary step in the automation process. Conversion of retrospective records, however, has always appeared to be such a formidable undertaking that few have been willing to face it. The case against retrospective conversion has been made by science and medical libraries on the valid grounds that most of the use of their collections is based on

Richard De Gennaro is Senior Associate University Librarian, Harvard University.

recent or current materials and that time will solve the problem. The same rationale does not hold true for general research libraries because the older materials in their collections are used more heavily and they continually acquire substantial quantities of retrospective materials. If the entire retrospective bibliographical record is not converted, these libraries will always be obliged to maintain their old manual systems along with their machine systems, and they will never get the maximum benefit from automation.

Libraries cannot seriously begin to design and implement "total" or "integrated" systems until they come to grips with and solve the problem of converting their retrospective catalogs into a machine-readable data base. This data base would become the foundation for subsystems for various operations, such as circulation control, searching, cataloging and catalog maintenance, and interlibrary loan service. It would also provide the means of generating the management statistics and information that are needed to improve library operations. Moreover, this comprehensive data base is obviously the foundation upon which networks must be built if the network concept is to become a reality. Actually, few librarians would question the desirability of having their entire catalogs in machine-readable form; they merely cannot believe that conversion can be accomplished, or accomplished at a reasonable cost. Let us discuss feasibility first and costs later.

Six years ago, no major research library had even begun in any serious manner to convert its retrospective catalog records to machine-readable form. There was no standard bibliographical format; the coding and printing of upper and lower case and diacritical marks was still poorly understood and difficult to accomplish with the available equipment. Almost no one had gained any significant experience in converting large files of complex bibliographical data and few librarians would have known how to use the products of such conversion if they had been available. Today the picture is drastically different. A considerable body of experience has been accumulated and a great deal of solid research and development has been done in the conversion of mass catalog files. Harvard's Widener Library shelflist conversion project was one of the early entries in the field. Routine conversion of its limited-entry shelflist in a local format was begun in 1965, and well over a half million entries of the estimated total of 1.6 million have been converted to date.[1] The Meyer undergraduate library at Stanford and the Ontario New Universities Library Project at Toronto, although more limited in scope, yielded

much valuable experience in conversion techniques. The Universities of Toronto, State University of New York-Buffalo, and Syracuse, among others, have converted portions of their shelflists or catalogs. The Library of Congress has developed the MARC II format and complex input systems, and has converted over 100,000 entries into machine-readable form in the two MARC formats. In England, the University of Newcastle-upon-Tyne has converted its entire catalog,[2] and Oxford University has embarked upon a major project to convert the pre-1920 Bodleian catalog using OCR (optical character reader) typewriters as the input medium and a format recognition program to lessen the manual editorial burden.[3]

The Institute of Library Research at Berkeley has done an outstanding job of research, development, and publication on the problems and techniques of mass bibliographic file conversion, principally in connection with its project to develop the design and specifications for a Technical Processing Center for the California State Library.[4] Its publications set a standard rarely equalled in the library automation field. In 1970 it embarked on one of the most ambitious, well-planned, and technically complex conversion projects that has been attempted to date, i.e., the conversion into MARC II structure and subsequent publication in book form of the estimated 900,000 records that form the 1963-67 supplement to the University of California's printed catalog. The project will use OCR-font typewriters and a highly-developed automatic field recognition system to facilitate input and minimize manual tagging and editing. The completion of this project in 1971 will mark the beginning of a new era in file conversion and the experience gained should be of considerable value to LC's Project RECON and other conversion projects.

The body of experience and knowledge gained in all these projects, together with the many improvements and developments in hardware and software that have taken place in the last few years, clearly indicate that the state of the art is now sufficiently advanced to support the large-scale conversion of complete bibliographical entries in the MARC II format.

While experience has shown mass conversion to be technically feasible at this time, it has also demonstrated that the cost is extremely high—in a range of one to two dollars per entry. Input keyboarding is only one of the costs and by no means the major one. Tagging the elements and editing the copy require the greatest effort and are the most difficult to accomplish since they demand personnel with train-

ing and experience in bibliographical work, and such persons are in extremely short supply. Computer and other machine costs are also significant, as well as project direction, administration, space, and other overhead costs. Another major category of expense which is frequently overlooked or underestimated is the very high cost of software development—systems design, programming, and program maintenance. Expense is not the only problem; it is difficult to find and hold the highly-skilled persons who are needed to do this complex technical work.

The issue is no longer whether the retrospective record can or should be converted, but rather how it should be converted and at what cost. Here we come to a critical point. Unless some over-all national plan for centralized conversion of a standard record in a standard format is developed and implemented in the near future, many libraries will begin (as many have already begun) to convert their own catalogs on an individual basis. The result will be the repetitive creation of expensive local conversion systems producing non-standard or sub-standard machine-readable entries. The combined cost of these separate efforts will exceed substantially the cost of a single centralized conversion effort which would provide a common bibliographical data base in the standard MARC II format from which libraries could draw a significant percentage of their catalog entries.

The RECON Working Task Force under the chairmanship of Henriette D. Avram has recently completed and published a comprehensive study entitled, *Conversion of Retrospective Catalog Records to Machine-Readable Form; A Study of the Feasibility of a National Bibliographic Service*[5] (hereafter referred to as the "RECON Study"). This excellent study, which was underwritten by the Council on Library Resources, Inc., has in one stroke raised the prospect of a national centralized retrospective conversion effort from the discussion and speculation stage to a level of systematic analysis and concrete planning. Since the RECON Study is now the basic document on LC-based conversion, this paper will necessarily draw upon and summarize many of its ideas and conclusions. For the serious student of retrospective conversion, library automation, or bibliographic networks, no summary can take the place of the full text of that report.

Most of the arguments that can be made in support of LC's centralized cataloging and card distribution service apply equally well to centralized record conversion. Indeed, if the MARC distribution service is a logical extension of LC's current card distribution service,

then creation and distribution of retrospective catalog data in machine-readable form is an equally logical extension of that service as well as of the MARC service itself. Conversion of its retrospective record is essential for LC's own future internal automation as well as for the card distribution service. Most of the experience, development work, and computer software that has been created for the MARC service is directly applicable to an LC-based retrospective conversion project. In short, conversion of LC records in the MARC format by LC is clearly the most reasonable and economical course to pursue; this is the major conclusion of the authors of the RECON Study.

So far, we have argued that retrospective conversion is necessary, that it can be done, and that it should be done centrally at LC with LC records as the starting point. Three other major considerations remain to be discussed: 1) the catalog or data base at LC which would be the most appropriate, 2) the principal technical and cost considerations, and 3) the over-all method of implementing the project.

With regard to the selection of the catalog or master data base to be converted, the RECON Study cites three important factors to be considered. There should be a high rate of duplication between materials covered by the data base selected and the collections of other libraries. The entries should have a high degree of accuracy and completeness, and certain types of entries should be excluded, such as serials and non-book materials. With these and other factors in mind there are only four catalogs which can be seriously considered for conversion: 1) the National Union Catalog, (NUC), 2) the LC shelflist, 3) the LC official catalog, and 4) the LC card division record set (a catalog of printed cards in LC card number and, therefore, roughly chronological sequence).

The NUC seems at first glance to be a likely candidate because of its size and comprehensiveness. In addition to its four million LC records, it contains seven million records that represent analytics, dissertations, local publications, foreign-language titles, etc., which are not on LC cards. These, however, are titles which are not held by many libraries and these entries do not come up to the standards of accuracy and completeness that are desirable in a master data base. Therefore the NUC was eliminated from further consideration.

Drawing on experience gained from converting the Harvard shelflist, the LC shelflist was this author's candidate for conversion in an article in *College & Research Libraries* published in 1967.[6] There would be many advantages to approaching conversion through the

LC shelflist if the shelflist were a reasonably accurate, up-to-date, and legible record. When the RECON Working Task Force considered the LC shelflist, it found that this file "contains a mixture of temporary, incomplete, and printed records with essentially no corrective changes beyond revision or updating LC class and book number. Nor are the cards legible enough to be microfilmed to provide a readable guide to locating the master records in the Official Catalog." [7] Because of these deficiencies the shelflist was eliminated from further consideration.

From the point of view of up-to-dateness, completeness, and accuracy, the LC official catalog would appear to be the most desirable candidate for a master data base. However, there are serious difficulties in using it directly for this purpose. The name portion of the catalog contains some twelve million cards of all kinds, and the task of searching out the four million discrete records produced since 1898 would be formidable. These records frequently contain numerous additions and corrections and would be difficult to use as a source document for first conversion. For these reasons the RECON Working Task Force recommends using the LC card division record set for first conversion and then bringing up to date the resulting record after comparing it with the master entry in the official catalog.

The card division record set consists of a master copy of the latest revised reprint of every LC printed card, arranged by card series and, within each series, by card number. The chronological nature of this catalog, its subdivision by series, and its legibility are potent arguments in favor of making it the starting point for conversion. The chief disadvantage is that not all changes in a catalog entry are cause for reprinting and therefore this record will have to be searched, compared with the official catalog entry and corrected to insure the level of accuracy and quality that a machine-readable data base of this nature requires.

Even with this disadvantage, the record set is still the prime candidate for first conversion because, like the shelflist, it has one of the characteristics most essential in a data base for a mass catalog conversion project: it is a large file that can be divided into a series of significant subsets that can be tackled and completed singly and used effectively as they are completed. The importance of this feature should not be underestimated. The Working Task Force recommends that the record set be divided by language categories (a tedious manual process) and that these categories be divided by time spans

of card series according to the table below.[8] This rearranged file would be microfilmed and a copy produced for the project; the original file would then be reconstructed for the card division.

Category	*Time span of card series*	*Number of records*
1. English language	1960-March 1969	386,000
2. Romance and German languages	1960-June 1970	381,000
3. English language	1898-1959	1,728,000
4. Other Roman alphabet languages	1960-June 1971	137,000
Nonbook materials	1960-June 1971	157,000
5. Slavic languages	1960-June 1972	225,000
6. Other non-Roman alphabet languages	1960-June 1973	256,000
7. Romance and German languages	1898-1959	698,000
8. All remaining catalog records	1898-1959	682,000

This table clearly demonstrates the advantage of being able to divide a large file into a series of significant segments to which priorities based on various considerations can be assigned. Such a strategy is a reverse chronological conversion sequence with priority assigned to the categories in greatest demand (and with the fewest problems). It will tie in nicely with the recommended rapid phasing-in of additional categories of current catalog data to be produced by the MARC distribution service. In no language category would retrospective conversion begin until the current records in that category were being produced by the MARC distribution service.

The RECON Study recommends that an initial conversion effort be made with English-language monograph records issued from 1960 to the beginning date of the current MARC service. This would be followed by conversion of Romance- and German-language monographs issued from 1960 to their beginning date in the MARC service (projected for June 1970). Both should be completed within four years. The third category would be English-language monograph records issued from 1898-1959. The conversion of other categories might follow the sequence of the table above or might be modified in the light of experience gained with the first three.

One of the very difficult and important problems that the study recognizes, but rightly reserves for further investigation, is how best to obtain standardized bibliographical records for items that are not now in the LC record set. In other words, how can the master data base be expanded to become a truly national union catalog or data base in machine-readable form? Just as the problem of retrospective conversion had to await certain hardware developments, the establishment of a standardized format, and the accumulation of some practical experience with conversion, so the larger problem of how to create and maintain a true national union catalog in machine-readable form must await additional hardware developments. It must also await the experience and knowledge that will be gained by the conversion, organization, and manipulation of a substantial body of retrospective and current records. The Working Task Force wisely recognizes this and recommends moving rapidly toward the design and implementation of a conversion system capable of handling a non-trivial pilot project of English-language records from 1960-68.

Retrospective conversion is no longer a technical problem awaiting the development of better keyboard input equipment or even the long-heralded advent of direct-read optical character recognition (OCR) equipment. The last few years have seen the development of several input keying devices including the magnetic tape inscriber and the OCR-font typewriter which are well suited to the mass conversion of bibliographic data. One of the most surprising findings of the RECON Study is that the cost of conversion by direct-read OCR equipment, when it is perfected, would be slightly more than conversion of unedited records by magnetic tape inscriber when all systems costs are considered.[9] The reason, according to the RECON Study, is that these devices will not be capable of reading non-Roman characters, diacritical marks, and other special characters. The machine would have to be programmed to reject records with an excessive number of unreadable characters and these records would have to be manually keyboarded. It was estimated that the number of records rejected might be as high as 10 percent. The cost of keyboarding corrections and unreadable records, added to the relatively high estimated cost of the OCR equipment itself, makes this alternative much less attractive than one might think. In any case, the importance of input keyboarding and the selection of input devices have been given more attention in the past than they deserved. These two factors together account for only 16 to 20 percent of the total unit conversion

cost in the LC environment and whether one device is slightly more efficient than another is a relatively minor matter. The selection of an input device in a conversion project may well be made on the basis of criteria other than cost.

Thus the need to await further breakthroughs on input equipment before undertaking large-scale conversion of bibliographical records has disappeared. We have the hardware for conversion, and we either have or know how to create the necessary software. The cost per record will probably never be much less than it is now, since 85 to 95 percent of the costs can be categorized as manpower, and only 5 to 15 percent as machine costs, and it is a fact that manpower costs are rising and machine costs are falling. It should be stressed that these cost ratios apply to the LC environment where the data has to be edited and corrected to the highest standards possible in order that it be acceptable for a national data base. In local environments where less complex procedures could be adopted and lower standards of accuracy could be tolerated, the total unit conversion cost might be significantly less and the manpower-to-machine cost ratio might be more nearly even. Indeed, a greater utilization of automatic format recognition programs might yield a more favorable result even in the LC environment. In any event, it appears that the chief obstacles to conversion are no longer technical; they are financial, political, and managerial.

While detailed consideration of technical and cost factors is not appropriate for this general overview, a brief review of these factors is essential for a basic understanding of how the records can be converted and at what cost. The RECON Study considered six input devices: 1) keypunch, 2) paper tape typewriter, 3) magnetic type inscriber, 4) on-line typewriter, 5) OCR-font typewriter, and 6) direct-read OCR (still under development). The keypunch and paper tape typewriter were eliminated as being technically unsuitable. The on-line typewriter and OCR-font typewriter were eliminated after a cost analysis showed them to be more expensive than the magnetic tape inscriber and direct-read OCR. The magnetic tape inscriber was deemed to be the most appropriate and least expensive device now practicable, while the direct-read OCR, although it is not fully developed, was retained as a possibility on the assumption that it might be used for some portion of the file when it is perfected. The elimination of the OCR-font typewriter as being too expensive is probably justified in the LC environment, but there is considerable evidence to

suggest that it may well be the most appropriate input device in other environments.

The manpower and machine unit costs of twenty technical alternative methods were analyzed and the four best ones were selected for detailed consideration. They were: 1) direct-read OCR (assuming its perfection in a few years) using a format recognition program, 2) unedited copy using a tape inscriber and a format recognition program, 3) partially-edited copy using a tape inscriber and a format recognition program, and 4) fully-edited copy using a tape inscriber. The resulting copy would in all cases be manually compared against the LC official catalog and corrected. The total cost per entry of the four alternatives ranged from a high of $1.87 to a low of $1.51 in the third alternative with 94 percent of the cost ascribable to manpower and 6 percent to machine costs.[10] Of the manpower costs in this alternative, $0.52 is for partial editing which in this context includes partial coding prior to input, post-editing to correct and augment the output of the format recognition programs, and editing of new data derived from comparing the interim records against the LC official catalog. This study confirms the conclusion that human editing in cataloging conversion projects is one of the most important cost factors and the trained personnel required are in short supply. That fact accounts for the intense interest that has developed in writing and utilizing automatic format recognition programs in such centers of conversion as Oxford, Berkeley, and the Library of Congress.

The format recognition program envisioned by the RECON Study analyzes the data in a partially pre-edited machine-readable record and automatically assigns tags or content designators and coded information which make explicit what is implicit in the textual information on the catalog card. Partial editing means that the records have been pre-processed by a human editor who has supplied some cues which increase the accuracy and reliability of the format recognition program. The utilization of these techniques reduces significantly the cost and difficulty of the conversion process by putting the burden of tagging and coding on the machine where it belongs. The Bodleian Pre-1920 Catalogue Project is successfully using a format recognition program that was initially written by John Jolliffe for the British Museum general catalogue.[11] The Institute of Library Research has successfully tested the ILR Automatic Field Recognition System, on several significant samples. The goal of the System "is to achieve a full MARC II record without any pre- or post-editing/tagging. The

computer recognition algorithms work with the existing format of the catalog card, and have no special input requirements." [12] The system was developed for use in the University of California Union Catalog Supplement Project and promises to reduce substantially the amount of human editing necessary and thus reduce the time and the cost of the project. This approach, if it proves feasible in actual operations, could be used for inputting current as well as retrospective records and might even be adaptable to the LC environment.

Applying the unit cost of the least expensive RECON Study conversion method, i.e., $1.51, the 386,000 English-language records in the record set from 1960 to March 1969 could be converted for an estimated $581,000. The cost of converting the 1,728,000 English-language records from 1898-1959 would be $2,602,000. To convert the estimated total of 2,114,000 English-language records would cost nearly $3,200,000. Since this is approximately half of the entire LC record set, the cost of converting the whole set would be on the order of $7 million.

The cost of systems design and software for a conversion system is estimated at $569,000 and is constant regardless of the number of records to be converted. The cost of hardware is based on the total number of records to be converted over a period of years and is therefore an extremely complex factor. However, for purposes of this discussion, the conversion, storage, and manipulation of the four million entries in the record set would require a two-shift computer system costing an estimated $7 million over an eight-year period. This system would support more than mere conversion operations; it would provide equipment for a national bibliographical service.

No matter how it is viewed, the total cost of retrospective conversion, including the cost of the systems design and software to accomplish it and the cost of the hardware necessary to support a national bibliographical service, is formidable. However, as has been said earlier, these costs seem far more reasonable and acceptable when viewed against the alternative, which is for each library to attempt to do its own conversion, a course which would produce a rich profusion of non-standard and incompatible records and systems at an enormous aggregate cost.

Another element in the cost picture which deserves more emphasis than is given by the RECON Study is the fact that full-scale mechanization of the card distribution service, upon which LC has embarked, is dependent upon conversion of major portions of the retrospective record. There is no doubt that this dual use of the machine-readable

data base along with dual use of the hardware and software would make the costs a good deal more acceptable. It should also be emphasized that a machine-readable data base will be a valuable property. It can be used to produce many kinds of marketable services and products which could contribute significantly to the support of the initial conversion as well as to the maintenance of the bibliographical system that will be based on it.

The Working Task Force recommended that the MARC distribution service be expanded as rapidly as possible to include all current cataloging done by LC in order to arrest, or at least slow, the growth of the retrospective record. It suggested that the cost of this expansion, along with some of the cost of retrospective conversion, might be budgeted as part of LC's regular operations, supplemented by grants and transferred funds. The research and development costs might well come from grants from private and governmental agencies with an interest in libraries.

Fortunately, planning for retrospective conversion did not end with the completion and publication of the RECON Study. Continuing the momentum that had been generated, the Library of Congress applied for and received an Officer's Grant of $25,000 from the Council on Library Resources, Inc., to implement the first phase of a RECON Pilot Project. This grant was made to convert the 85,000 English-language monograph titles cataloged during 1968 and those English-language titles cataloged in 1969 but not included in the MARC distribution service.[13] The conversion will provide a practical situation to test and study the various conversion techniques as well as the concepts and techniques of partial-editing and format recognition as outlined in the RECON Study so that the best methods for future conversion efforts can be determined. In addition, a representative sample of from five to ten thousand older titles in English and other Roman-alphabet languages will be drawn from the record set for further detailed analysis and testing.

Thus, the important task of retrospective conversion is moving from the study stage to active experimentation in the form of a pilot project. Experience with library automation has shown that this method of proceeding in stages is the one most likely to be successful in accomplishing a difficult and complex task. The development of the MARC II format and distribution service followed a similar pattern with excellent results.

The RECON Study leaves many technical, organizational, and pro-

cedural questions unanswered. How can such a massive file be organized and maintained efficiently and effectively? How can the data base be expanded into a true national bibliographical system with locations? How can other libraries draw entries from the data base for their own use and add their unique holdings to it? How will serial entries be handled? In what forms and on what financial basis will the data be distributed to libraries as well as to firms desiring to exploit its commercial possibilities?

Some of these questions will be answered by the first phase of the RECON Pilot Project; others will be answered only in later phases. In any case, despite the numerous problems and islands of ignorance that remain, a significant beginning has been made on the task of converting LC's retrospective bibliographical file to machine-readable form, and cautious optimism is in order.

References

1. De Gennaro, Richard. "Automation in the Harvard College Library," *Harvard Library Bulletin,* 16:217-36, July 1968.

2. University of Newcastle-upon-Tyne. "Catalogue Computerisation Project" (Interim Report, Sept. 1, 1967 to Aug. 31, 1968). Sept. 1968.

3. Brown, Peter. "The Bodleian Catalogue as Machine Readable Records," *Program: News of Computers in Libraries,* 3:66-69, July 1969.

4. Cartwright, Kelly L. and Shoffner, Ralph M. *Catalogs in Book Form; A Research Study of their Implications for the California State Library and the California Union Catalog, with a Design for their Implementation.* Institute of Library Research, University of California, Berkeley, 1967; Cunningham, Jay L., *et al. A Study of the Organization and Search of Bibliographic Holdings Records in On-Line Computer Systems: Phase I* (Final Report Project No. 7-1083, Grant No. OEG-1-7-071083-5068). Institute of Library Research, University of California, Berkeley, March 1969. See especially Chapter IV: "Data Base Development," pp. 75-116; and Sherman, Don and Shoffner, Ralph M. *California State Library: Processing Center Design and Specifications.* 3 Vols. Institute of Library Research, University of California, Berkeley, April 1969.

5. *Conversion of Retrospective Catalog Records to Machine-Readable Form; A Study of the Feasibility of a National Bibliographic Service.* Prepared by the RECON Working Task Force, Henriette D. Avram, Chairman, John C. Rather, ed. Washington, D.C., Library of Congress, 1969.

6. De Gennaro, Richard. "A Strategy for the Conversion of Research Library Catalogs to Machine Readable Form," *College & Research Libraries,* 28:253-57, July 1967.

7. *Conversion of Retrospective Catalog Records . . ., op. cit.,* p. 25.

8. *Ibid.,* p. 31.

9. *Ibid.,* p. 97.

10. *Ibid.,* p. 99.

11. Joliffe, John. "The Tactics of Converting a Catalogue to Machine-Readable Form," *Journal of Documentation,* 24:149-58, Sept. 1968.

12. Sherman, Don. "Initial Progress Report on Automatic Field Recognition" (Technical Paper No. 5). Institute of Library Research, University of California, Berkeley. Aug. 18, 1969, p. 1.

13. U.S. Library of Congress. *Information Bulletin,* 28:427-28, Aug. 21, 1969.

Methods and Research for Design of Information Networks

VLADIMIR SLAMECKA

THIS ARTICLE CONCERNS questions of research appropriate to the design of networks of library and information services and is organized into four sections. The first section reviews a selection of previous statements on research areas held relevant and important to the design of library and information networks and suggests a derived functional framework into which most of these viewpoints can be accommodated. The second section briefly distinguishes between two basic goals of information network design, and it characterizes their respective design approaches in terms of methodology. In the third part, information networks are postulated which assume the accepted present-day premises of library and information service objectives, and the research effort conducive to the design of such networks is discussed. The last section of the paper attempts to indicate major assignments for research in the design of information networks which extend the scope of current information services and objectives.

With varying degrees of conviction and emphasis, research on information service systems and networks has usually been viewed as a necessary or desirable component of efforts, the goal of which is the design of such systems and networks. If the role and topical direction of such research were to be summarized in a single sentence, it would be difficult to do so more concisely or encompassingly than does the statement made in 1960 in the U.S. Senate: "The most pressing need therefore is for the development of reliable methods for studying and assessing requirements, for determining the role of information and information services in science, and for measuring the value of information and the utility and effectiveness of present and proposed services." [1]

Vladimir Slamecka is Professor and Director, School of Information and Computer Science, Georgia Institute of Technology, Atlanta.

The following sample of comments on the role and the areas of research relative to the design of library and information service systems and networks represents viewpoints which are consensuses rather than opinions of individuals. These comments demonstrate varying degrees of concern by the profession and by the scientific community, and they impart an overview of problem areas and of approaches to solutions which these groups have considered significant. (No attempt is made in this paper to review or summarize recent and current research on issues relevant to information networks; the reader is referred to several publications [2] which accomplish this task quite well.) Two types of comments are especially indicative: those emanating from reflective deliberations of qualified study groups and committees, and those underlying large-scale experimental efforts in the direction of modern information-sharing networks.

Project Intrex, conceived in 1965, partially qualifies among the latter efforts. Although planned primarily as a program of experiments addressed to the broad problem of access to bibliographic materials and conducted in an atmosphere of a model library, it proposed to "explore a range of ideas designed to promote the integration of university libraries into the national (and, ultimately, international) network of information centers." [3] The suggested direction was to develop a direct, inexpensive interface with any citation retrieval service, such as that of the National Library of Medicine or the National Aeronautics and Space Administration, and to be available to the M.I.T. community as well as to a variety of other users. A supporting program of research mentioned by the Intrex planners was to attack "the primitive state of two critical items: [computer] consoles and interaction languages." [3]

A study conducted by the Interuniversity Communications Council (EDUCOM) in the summer of 1966, began to explore the establishment and operation of information networks with an enumeration of a few key problems or problem areas, including those of formal organization, financial support, standardization and quality control, copyright, evaluation of operation, and the roles of the individual and the community. The study produced a technical plan [4] for the establishment and evaluation of a pilot national information network, implying that the state of the art had sufficiently progressed to permit its design and experimental operation. However, another EDUCOM conference, held in the fall of 1967, discussed the need for more accurate data relating to the decision on the storage medium for textual

portions of library collections; [5] subsequently, EDUCOM announced it would study numerous different characteristics of physical information records prior to considering the design of experimental equipment and conducting experiments, tests, or studies.[6]

The Bolt, Beranek and Newman study for the Council on Library Resources, Inc., on concepts and problems of future libraries [7] urged advancements of relevant sectors of technology through the mechanism of positive interaction among the disciplines of library science, computer science, systems science, and the behavioral and social sciences. Concerned primarily with the incorporation into information services of procognitive functions, the report identified the following needed research efforts: the determination of basic characteristics of the relevant network that interrelates the elements of the fund of knowledge (that is to say, the development of an effective, formal analytical semantics); the development of advanced memory systems likely to come from studies of higher order associative memories; the development of fast processors for manipulating complex information representations and structures; the development of advanced displays and control for man/computer interaction; the development of procedure-oriented, field-oriented, and user-oriented languages with which to control the processing and application of the body of knowledge; an understanding of machine processing of natural languages; and the development of multiple-access computer systems.

The Carter report provides a brief review of the feelings on research by those groups and individuals who concerned themselves with the issues and plans of national document systems. Summarizing the discussions of further needed research in fifteen such plans offered up to 1967, the report states:

> The most frequent recommendation was to develop "standards" of one kind or another (e.g., standards for indexing, abstracting, communications, hardware, cataloging). A notable feature of these recommendations was that the plans called for *what* was to be standardized, but made no recommendations about *how* this was to be accomplished. The next most frequently mentioned areas for additional research were those of developing educational curricula and of studying the needs for better equipment and facilities. Other important topics proposed for research and investigation included: (1) Informal communications. (2) User needs and characteristics. (3) Inventories of libraries and other information resources. (4) Improvement of publications and technical writing. (5) Copyright and patent problems. (6) Classified and proprietary information.[8]

The Carter report itself was concerned primarily with organizational issues; it did, however, recognize research as an important function of the "capping agency" (which would manage the design and operation of the national document system), and called in its basic propositions for a "balanced program that will give appropriate weight to all aspects of the necessary research and development." It proceeded to single out the following among the areas of most pressing current needs for research: user studies to determine both the recognized and the unrecognized needs of users; document representations by both manual and automated means; evaluation tools and techniques for assessing the adequacy of information systems; communication of information to clarify the role of documentation and other media of information transfer; and equipment for the storage, manipulation and transmission of information to users.

The most recent consideration of research needs relative to the pressing national problem of scientific and technical communication was offered by the Committee on Scientific and Technical Communications (SATCOM) of the National Academy of Sciences—National Academy of Engineering. (Of the previous policy milestones in science communication, the 1963 report of the President's Science Advisory Committee, the Weinberg Report, avoided specific recommendations on needed research areas, aside from urging all organizations concerned with technical information to investigate the new techniques and ideas and giving its approval to "practical tests of new modes of technical communication." [9] In the 1965 report on a national information network in science and technology issued by the Committee on Scientific and Technical Information,[10] COSATI was particularly concerned with indexing and abstracting costs and effectiveness.) The SATCOM study,[11] which concerns itself with many of the issues of the Weinberg Report, repeatedly harped on the lack of understanding concerning communications research and of data on various systems aspects of the process of scientific and technical communication. As for desirable areas of research, SATCOM recommended several areas of study, research and experiments, including the following: comprehensive analyses and experiments on the functioning of different parts of the scientific and technical communication network and on its overall operation, the development of measures of the value of information services and ways of overcoming user apathy or resistance, experiments involving the use of the computer in conjunction with human workers for the preparation of indexes, the development and evalua-

tion of languages for describing the formats of files and other types of digital communication systems, and the development of standard structures for each widely used bibliographic documentary information element. SATCOM further urged scientific and technical societies to participate in research, recommended large-scale experiments, and suggested that the federal government establish a single group to plan a unified program of critical experiments of operational scale.

It would be unfair to expect from any of these well motivated and knowledgeable quarters a definitive discussion of research requirements: some were concerned with narrower, usually political, aspects of information networks, and a few with the broader problems of scientific and technical communication. In addition to these eclectic consensuses, numerous personal opinions on research stand recorded by individuals from diverse walks of life. Predictably, they are uneven in both purpose and content, ranging all over the research map (and even rejecting it as either irrelevant or resolved). Jointly, however, these discussions strongly imply that there has been no attempt made so far to provide a reasonably detailed outline of the social, technical and management problems whose solution is necessary and sufficient for the design of advanced information networks.

To understand the reason for the absence of a research plan and schedule, or even of a research manifesto, it is helpful to consider the dual form in which research can exist. These two forms are perhaps best illustrated by an analogy which contrasts the self-directed, sporadic and spontaneous research in the aerospace and aerospace materials sciences before 1960 and the defined, planned research effort in these sciences under the Apollo program of NASA. The distinction lies, of course, in the degree of definition of the research purpose and mission. The absence of a comprehensive research schedule for the design of advanced information networks thus infers a lack of technical plans to implement such networks.

What is the range of the research areas which bear on the design of library and information networks? Table 1 attempts to suggest a framework for placing social, technical and administrative considerations into a functional, discipline-oriented relationship. While nearly all calls for research fall into this framework, there is less agreement on the relative value of individual areas and topics.

A *plan* to design and implement an advanced network of library and information services is actually a statement expressing known relations and dependencies of significant component activities and pro-

cesses of the task. Once such a plan has been drawn up, it is relatively easy to identify what problems, if any, should or must be researched. In the absence of a plan, however, discussions of proposed or necessary research should at least outline the concept or form of the information network to which they refer; without such an outline, the question of research need and value often remains open.

For the purpose of our discussion, a generalized concept of library and information networks may be given in terms which are reasonably familiar. As its basis, we assume the existence of a physical network of information processing devices communicating in a machine-to-machine mode. This network forms a physical substratum which provides a switching capability permitting people and machines to access a variety of partially compatible information and data banks. The content of these banks then can be maintained and manipulated by means of instructions stored in programs, or conversationally from local or distant inquiry stations, to provide a range of information products and services. Since commercial as well as military data communications networks are under advanced development in the United States, mechanized library and information networks are undoubtedly technically feasible.

Table 1. Research Areas Relating to Design of Networks for Library and Information Services

A. *SOCIAL SCIENCES*

1. Environment
   - Functions and relations of component agencies in the network
   - Identification of new services
   - Political, legislative and legal considerations
   - Effects on man, society, science
2. Market
   - Identification and description
   - Information requirements and uses
3. Manpower
   - Personnel requirements and characteristics
   - Education and training
4. Management
   - Organization
   - Management: planning, operations, control

B. INFORMATION SCIENCES

1. Theory of Information
   - Semiotics (syntax, semantics, pragmatics of natural and artificial languages)
   - Information processes (generation, collection, coding, organization, transmission, transformation, storage, use)
   - Information measures
2. Human Engineering
   - Man-machine communication
   - Man as information processor
3. Information Processing Technology
   - Hardware design and operating characteristics
   - Software languages, systems
   - Communications engineering
4. Information Systems Engineering
   - Information systems properties (structure, behavior)
   - Information systems analysis and synthesis
   - Methodologies of complex systems design evaluation
   - Economics of information systems and networks
   - Management of design and operations

The development of advanced networks of library and information services can proceed via two principal approaches. If services and systems of services already exist whose objectives and functions are acceptable, the development of a network amounts to a transition from the present-day state of these services to a higher level. The primary goal of this development is then an improvement in the *efficiency* of the system processes and performance.

The extreme of the alternate approach begins with a formulation of new objectives and functions for a system yet to be designed. The designer follows a procedure which embodies rigorous elements of the scientific method. For systems of information services, this procedure contains the following sequence of three phases:

1) Assessment of the market
   - definition of market
   - identification of information uses
2) Design of information services
   - standard products
   - special services

3) Design of the information system
   information store
   process and operations design
   quality control

In contrast, the redesign of systems of services which retain their basic objectives and functions involves the second and more often, only the third phase.

Using this distinction of approach, we can conveniently identify and distinguish between design efforts which propose (and are restricted to) the improvement of efficiency of existing information services, and those which seek new objectives as the basis for their services and systems. The thesis elaborated in this article is that the two goals and approaches impose different roles and challenges on research.

When reviewing the methodologies utilized in present-day ongoing designs of library and information networks, it becomes apparent that they are primarily concerned with redesigns of systems of existing services. While these redesign efforts frequently entail an upgrading of some parameters of existing services (e.g., speed, comprehensiveness, etc.), the consideration of new services and especially of new markets is conspicuously absent from ongoing activities in library and information networking.

It is not difficult to find an explanation for this phenomenon: there are already in existence complex systems consisting of bibliographic information banks, products and services, vendors, and of generators and users of information. It is also not difficult to find a justification for the redesign or continued development of these systems. Although some parameters of their services are admittedly suboptimal, their improvement appears feasible via applications of newer technologies and organizational approaches. The justification for further investment in these systems can be argued even in the face of the difficulty of assessing the temporal economic value of their services since they are a significant component of the process of codification of human knowledge in that they create an external memory for mankind to aid in the carrying out of numerous human functions and activities. The bibliographic function clearly belongs among those which appear very desirable for society to sustain.

The currently attempted and planned networking activities in the library and information professions have as their goal an introduction

of greater efficiencies in the existing and assumed functions of bibliographic control and services. These efficiencies are to be derived through the sharing or networking of various selected processes and of data in existing systems. The emerging networks will thus provide for centralized bibliographic record-keeping on a geographical and/or subject basis; cooperative acquisitioning and technical processing of materials; cooperative, coordinated production of various types of bibliographic aids; and optionally, for a capability of decentralized inquiry against compatible record files. While efficiency improvement of existing library and information services can be a desirable goal by itself, it is apparent that in accepting the objectives and premises of these services, the current networking designs are not concerned with fundamentally new approaches to improve the *effectiveness* of information communication in society.

Since it appears that for the next several years we shall be committed to a networking of existing library and information agencies and services, it is appropriate to ask what research is necessary or desirable to assist in this effort. The methods being used to induce efficiency are of two types: technical and organizational. The technical approach rests primarily in the mechanization of physical processes and their elements; the organizational method employs coordination, cooperation, and partial centralization. To the extent that the success of the current efforts at networking of services will be reflected by the efficiencies attained, the necessary or desirable research can be expected to relate principally to these two methods. The following few paragraphs briefly discuss indicated areas of research.

The frequently stated design desideratum [12] that the information service systems accommodate the information communication process contiguously, exposes that a serious lack of knowledge lies at the interface of two system components. There is at present a discontinuity between these two components—the physical store or inventory of information and the media of access or directories to the store—which prevents a smooth and convenient transition and interaction between directories and documents or full-text information. (The mechanization of directories tends to stress this discontinuity; whereas previously I had to walk from a card catalog into the stacks, the possibility of having a library catalog on microfiche in my office or on loan at home is a frustration like having a telephone book in a city without a telephone service.) In technical terms, the unresolved problems are the relation, roles, and the consequent development of the analog and

the digital technologies and their optimal symbiosis in a library and information network. There appears little doubt that for some time to come the two information forms, analog and digital, shall coexist in our information systems; there is a lack of clarity at the present time concerning the nature and boundaries of this relationship, and of the methods for viably tying both into our systems of services.

A second area of technical research concerns the characterization of the traffic of information in bibliographic networks. Data are lacking on the types of communication which such a system can anticipate and on the distributions of the traffic between generators, vendors and users over various parameters. The actual configuration of the networks is very much dependent on such assessments, as are reasonably accurate transmission requirements and the choice of transmission media (mail, cable, microwave). The doubts of the EDUCOM 1966 study on information networks as to the "ability to produce anything meaningful without firm specification to traffic loads" [13] remain warranted.

Another broad problem area is that of compatibility and standardization. Although we must constrain both the system operator and user by rules, there clearly exists some freedom in slanting these constraints one way or the other. It might be asserted that constraints on the system input are less dangerous, technically more desirable, and in practice easier to implement than a stringent stifling of the users' options and of their not-so-consistent behavior patterns. What is not so clear, however, are the necessary minimum levels of standardization for either the system or the users.

The basis of the problem of compatibility lies in the behavioral patterns of humans with respect to their uses of information, and more fundamentally, in the properties and in human use of natural language. Ultimately, the criteria for the standardization of all system elements must be derived from studies of these areas at research levels which are only now being unfolded (e.g., semiotics). Meanwhile, it is desirable to ascertain the dimension of the actual need to switch automatically from one information file to another; to investigate authoritatively the several methods through which file convertibility can be implemented within and among information networks; and to propose, as a minimum, the design specifications for a file-description and inquiry language or languages.

A second category of desirable research concerns questions of political and organizational nature. It is clear that major efficiencies in the

existing library and information systems can be realized at the level of cooperative networking. The types, sizes and characteristics of these networks are subject to an interplay of a variety of factors—political, economic, geographic, and human. Thoughtful analyses of the possible, desirable and permissible categories of networks and network nodes should yield another element of the data necessary for the systematic development of a national network.

A major threat of delay to the networking of library and information services appears to be the less-than-enthusiastic attitude of the decision-making echelons of the library profession. It is improbable that this attitude can be dissolved solely by standard management techniques which apply political pressures or through financial incentives which may encourage divisiveness. We must recognize the serious social overtones inherent in this resistance to networking and mechanization, as well as the symptom of a profession in crisis. It is perhaps here, in the area of social science research in librarianship, that a most urgent effort is vital in order to find new directions for a profession in transition.

It should be observed, concomitantly with the identification of several areas of propitious research, that the state of the art is sufficiently advanced to warrant the development of existing library and information systems in the direction of networking their services. Impressive information processing and communication devices are available; the techniques to store and manipulate information are given, although not in anything near an adequate supply; progress exists in the development of machine-readable resources; encouraging agreements are recorded on common formats of principal elements of the bibliographic inventory. The principal question to be asked and answered today then must be the following: What level of network design and implementation is presently justified?

In part, this is a technical question, and research in the areas described above should be formulated so as to provide the required answers. To the extent that the question is related to the possible levels of efficiency in existing information services, it is a question of cost and of return on the investment. To give meaningful answers, it is necessary to ascertain and make known the operating costs of existing library and information services; the design and operating costs of alternate designs of several logical levels of service networks; and a plausible assessment of the apparent desirability and value of the service improvements which such designs make available to users.

Given such cost analyses and continuing to accept the assumed premises of the social and economic value of our information services and of their objectives, the direction and level of further development of our library and information systems can be indicated with authority.

The incentive for this further development of information systems is attractive: the possibility of greater efficiencies in the inventory and distribution of information, in the maintenance of these repositories, and in the less tangible but positively favorable effects on information users.

The relative intangibility of the effects on information users has, however, lately raised questions concerning the justification of proposals to implement information networks other than those of cooperative libraries. These questions point to the lack of consistent evidence that networking of present-day information systems will improve the *effectiveness* of their services. More fundamentally, they question implicitly some of the premises which underlie the objectives of existing information services in that the objectives appear to be based on a "lack of knowledge and objective data about requirements, about the parameters of the problem, about the value of existing services and systems, about the effectiveness of these systems, about suitable techniques for measurement and evaluation, and about suitable criteria of effectiveness and value." [14] The result is an impasse; while it seems probable that the physical communications platform being developed in the United States is technically capable of sustaining a nationwide network of library and information services, the would-be designers of this network are finding it difficult to justify its implementation.

The inquiry into the objectives of present-day library and information systems and into the bases of their future development is founded on one searching issue—their value. The concern is a proper one since all of the parties involved—the users, the funders, and the designers—must desire that the further large-scale development of our systems of services be justifiable not only in terms of operating efficiency but also in terms of the effectiveness with which it meets its objectives. Apparently the value parameter of our design objectives must ring more convincingly. (The faint hope that the evidence of the low exploitation of library and information materials and services is perhaps inadvertently in error, or that better techniques of system evaluation might reverse the negative conclusions, are not well founded.)

Indeed, it appears that possibly the only way to proceed is to direct serious and conscientious research toward a category of design para-

meters which until now have been assumed or ignored, i.e., the definition of information markets, and the description of the information-using processes in these markets. The measure of effectiveness and the design of effective information services and of their nationwide network are predicated on this research mission of primary significance.

The identification and description of markets is an early step in the design of any information service; the objectives and the functions of information systems are both dependent on an assessment of the target markets for their services. To be successful in the present-day climate which asserts that information is a resource having a significant utility potential, the information services should be directed at markets whose characteristics include at least one of the following: high volume or frequency of use, or high value of information used.

Apart from the library cataloging and inventory-keeping functions, virtually all present-day systems of bibliographic services have a single market: the community concerned with scholarship, research, development, and production of either knowledge or goods. These systems dispense scientific and technical information to the scientific and technical community, whether it be found in the university, government, industry, or the professions.

The design of information services by the library and documentation professions is predicated on the concept and model of scientific and technical communication as a process of continued information metamorphosis by a Gargantuan organism—the scientific and industrial establishment—which thrives in proportion to its ability to digest its own product, information. It is logical to infer from this model that "new science and technology rest firmly on the base of information generated in the past," [15] and easy to accept as axiomatic the conclusion that the need for information services in this establishment is one of crucial importance and enormity of proportion. To the dismay of nearly everyone, however, this axiom has been found most difficult to uphold whenever it was put to the test. Much of the evidence produced by our professional studies of "information use" is actually contradictory to it. Why this perplexity?

A British librarian, W. L. Saunders, recently made this observation:

> Much of our professional activity in university libraries is based on certain assumptions about the role and the significance of information and libraries in one very important branch of creative activity

—namely academic research and scholarship. It could well be that library services and "information" in the librarian's sense make a very much smaller contribution to the creative process than we librarians like to believe.[16]

Let us suggest why this could well be. Knowledge, or information, exists in two "memories"—one internal to the human mind, the other external to it. Man's interaction with information can draw on both of these banks. The crude evidence we have about this interaction suggests that the researcher or engineer—in short, the creative man—draws very heavily on his own information bank (memory) and, often, on the internal information banks of the visible or invisible college of which he is a member. If we reflect that he automatically stores the information he produces in his internal memory, and that his subject areas of interest or activity remain largely contiguous, it indeed becomes quite plausible that our creative man needs to resort to the external information bank only seldom, and that his need for our present forms of information services is far less frequent than we like to believe. The confusion between the reasonable axiom and the evidence contradicting it is one between the need for and the use of information (which *is* there, voluminous and frequent) and the need for and the use of extra-memory information services (which is there only at times).

There are other factors at play, of course, such as motivation. However, if we accept the evidence that the a priori expectation of an overwhelming, continuous need and necessity in the creative elements of society for present-day information services is somewhat mythical, significant implications apply to our past and current activities, and serious assignments arise for our impending directions of research.

A considerable research effort of the recent years has been motivated by the growing volume of disturbing data which suggest that the users of scientific and technical information are rather immune to both new and old information products and services. It was quickly accepted that the apparent low exploitation of these products and services must be a property of their design, to which assumption research responded vigorously with submolecular analyses of every imaginable variation on the format of information products and services, and on the techniques of their derivation, use and evaluation. The literature attests to hundreds of studies of cataloging, classification, indexing, abstracting, and retrieval theories, techniques and products, and to extensive investigations of precision and other measures of ap-

parent system performance and behavior. It is not intended here to deprecate the quality of the hundreds of man-years of research effort so invested, or to minimize the insights gained into information processing techniques, operating efficiency, and into the qualitative and quantitative relations among various system elements. At the same time, the inference is obvious that if the assumed characterization of the science information market incorrectly represented its process of information use, much of this research must be suspected of being irrelevant and redundant to the issue of information service effectiveness.

Without dispute, the scientific and industrial community is one of the markets for library and information services. Meanwhile, it is a very timely question to ask what other, perhaps more lucrative markets, for science information are waiting to be served. Almost intuitively, such a market must be education. The acquisition of knowledge in an organized, codified form is the objective of education, and it may be viewed as a process of information transfer between the external and internally memories of the learner. The learning process invites advanced information services and systems, and the technology-minded trend of development of this immense market has already brought it within our sphere of concern.

It is probable that other information markets of the magnitude of education are waiting for recognition. If the information-handling profession intends to become a service industry, market research and analysis are necessary to its planning and justification. The immediate and significant task for research is to formulate a program of perceptive studies to identify such information markets, and to analyze their characteristics and their potential as customers of advanced information services and systems.

The identification and description of information market characteristics must be accompanied by studies of the uses of information in these markets; information services and systems can be designed and operated successfully only if they meet well-defined requirements. That this has been recognized all along is attested to by the numerous "user studies" carried out over the past twenty-five years in nearly all parts of the world. Notwithstanding this effort, it is not inaccurate to observe that these studies have failed to define the information-using process of their subjects in a manner meaningful to the information system designer. The unwitting evidence of a discrepancy between presumed needs and actual use of scientific

information produced by these studies reduces, without ruling out, the possibility that existing or proposed information services can be made more effective (that is, that they can have a greater effect on the efficiency of scientific and technical work). What kind of research is then needed to provide better data on the information use requirements by our markets?

The awesome goal—to understand the working mechanism of the human mind—is obviously too formidable for a realistic research mission and unlikely to yield useful data in time to identify reliable design requirements for the next developmental phase of library and information systems. The desirable research level lies between this goal and the collecting of sporadic data on users information habits. Such a level is given, we believe, by a program of information-oriented studies into selected symbol manipulating processes of man. These functional processes are characteristic of our information markets: problem-solving of the research and development community, design of the engineering community, learning of education, decision making of management. The immediate objective of such a program of research is not so much elucidation of the functioning and behavior of the human mind (in the sense attempted by research in neurophysiology, bionics, etc.) as the intermediate sequential description of the functional components principally interacting in these processes: goals, methods, and data.

Of greatest immediate interest in our analysis of requirements for the design of effective information services is the role which external information memories can serve in the interaction of these components of human symbol-manipulating processes. Offhand, certain differences in the process characteristics of various information markets are obvious. For example, the information-based process of learning draws heavily on external information memories. In contrast, the interaction with external information stores may be expected to be less frequent, less linear and, perhaps, less predictable in the case of the problem-solving, creative activity in science. A comprehensive categorization of the elements of information in these processes according to various parameters (form, purpose, function, frequency, dependency, precision, value, and perhaps others) should permit an insight into both the requirements of user markets and the characteristics and structure of our information repositories. Founded on such knowledge, the subsequent phase—the design of information services and systems—will be

able to consider objectively questions on which today there exist only opinions.

The suggested direction of research is also logical and appropriate to the technically and politically difficult question of effectiveness. The present-day lack of satisfactory approaches to the measurement of effectiveness and value of information services is a partial consequence of the relative fallacy of past user studies; it is obvious that the assessment of effectiveness of any service is objectively impossible without the knowledge of the specific requirements which such service proffers to meet. Once these requirements have been assessed, however, the question of information service and system effectiveness becomes technical rather than speculative.

An incidental result of the requirements analysis research could very well be a partial insight into a most difficult property of information measure—its value. The development of a pragmatic measure of information utility must rank as a most important task of information science research in the next decade. An adequate functional description of the major information-using processes of man may, however, shift the responsibility for applying this measure on the information user, and away from the service designer and vendor. This is indeed desirable; the value of information is intrinsic to information processes and goals of users and society, and as such it must be only monitored, not determined, by the information industry.

References

1. U.S. Senate. Committee on Government Operations. *Documentation, Indexing, and Retrieval of Scientific Information* (86th Cong., 2d Sess., Senate Report No. 113). Washington, D.C., U.S.G.P.O., 1960, p. 110.

2. Henderson, Madeline M., *et al. Cooperation, Convertibility, and Compatibility among Information Systems: A Literature Review* (National Bureau of Standards, Miscellaneous Publication 276). Washington, D.C., U.S.G.P.O., 1966; and Cuadra, Carlos A., ed. *Annual Review of Information Science and Technology.* Vol. 3. Chicago, Encyclopaedia Britannica, 1968; and U.S. National Science Foundation. Office of Scientific Information. *Current Research and Development in Scientific Documentation.* Vol. 15. Washington, D.C., National Science Foundation, Office of Scientific Information, 1969.

3. Overhage, Carl F., and Harman, R. Joyce, eds. *Report of a Planning Conference on Information Transfer Experiments, September 3, 1965.* Cambridge, Mass., M.I.T. Press, 1965, pp. xvii-xviii.

4. Interuniversity Communications Council. *EDUNET; Report of a Summer Study on Information Networks, Conducted by the Interuniversity Communications Council (EDUCOM)* (Information Science Series). Written and edited by George W. Brown, *et al.* New York, Wiley, 1967.

5. "Document vs. Digital Storage of Textual Materials for Network Operations," *EDUCOM Bulletin*, 2:1-5, Dec. 1967.

6. "Library Digital-Storage Study," *EDUCOM Bulletin*, 3:15, Nov. 1968.

7. Licklider, J. C. R. *Libraries of the Future*. Cambridge, Mass., M.I.T. Press, 1965.

8. Carter, Launor F., *et al. National Document Handling Systems for Science and Technology* (Information Science Series). New York, Wiley, 1967.

9. U.S. President's Science Advisory Committee. *Science, Government, and Information*. . . . Washington, D.C., U.S.G.P.O., 1963, p. 30.

10. U.S. Federal Council for Science and Technology. Committee on Scientific and Technical Information. *Recommendations for National Documents Handling Systems in Science and Technology*. Washington, D.C., U.S. Dept. of Commerce, National Bureau of Standards, Institute for Applied Technology, 1965.

11. National Academy of Sciences. National Academy of Engineering. *Scientific and Technological Communication*. Washington, D.C., National Academy of Sciences, 1969.

12. U.S. Senate, *op. cit.*, p. 102.

13. Interuniversity Communications Council. *EDUNET* . . ., *op. cit.*, p. 282.

14. Henderson, Madeline M., *op. cit.*, p. 98.

15. National Academy of Sciences. . . . *Scientific and Technical Communication, op. cit.*, p. 1.

16. Saunders, W. L. "The Establishment of Priorities." In *Research into Library Services in Higher Education: Papers Presented at a Conference held at the University of London on Friday November 3, 1967*. London, Society for Research into Higher Education, 1968, p. 23.

Library Trends

VOLUME 18

July 1969 - April 1970

This Page Intentionally Left Blank

Library Trends

Index to Volume 18

PREPARED BY PEGGY PODLASEK

This Page Intentionally Left Blank

Complete List of Library Trends Issues in Print

		Title	Editor	Date
V. 1, N. 1		Current Trends in College and University Libraries	R. B. Downs	July 1952
1	2	Current Trends in Special Libraries	H. H. Henkle	Oct. 1952
1	3	Current Trends in School Libraries	Alice Lohrer	Jan. 1953
1	4	Current Trends in Public Libraries	Herbert Goldhor	April 1953
V. 2, N. 1		Current Trends in Libraries of the U. S. Government	Verner W. Clapp Scott Adams	July 1953
2	2	Current Trends in Cataloging and Classification	Maurice F. Tauber	Oct. 1953
2	3	Scientific Management in Libraries	Ralph R. Shaw	Jan. 1954
2	4	Availability of Library Research Materials	Dorothy M. Crosland William P. Kellam	April 1954
V. 3, N. 1		Current Trends in Personnel Administration	Bernard Van Horne	July 1954
3	2	Services to Readers	Leslie W. Dunlap	Oct. 1954
3	3	Library Associations in the United States and the British Commonwealth	David H. Clift	Jan. 1955
3	4	Current Acquisitions Trends in American Libraries	Robert Vosper	April 1955
V. 4, N. 1		Current Trends in National Libraries	David C. Mearns	July 1955
4	2	Special Materials and Services	Andrew H. Horn	Oct. 1955
4	3	Conservation of Library Materials	Maurice F. Tauber	Jan. 1956
4	4	State and Provincial Libraries in the United States and Canada	Paxton P. Price	April 1956
V. 5, N. 1		American Books Abroad	Dan Lacy Charles Bolte Peter S. Jennison	July 1956
5	2	Mechanization in Libraries	Arnold H. Trotier	Oct. 1956
5	3	Manuscripts and Archives	R. W. G. Vail	Jan. 1957
5	4	Rare Book Libraries and Collections	Howard H. Peckham	April 1957
V. 6, N. 1		Current Trends in Circulation Services	Wayne S. Yenawine	July 1957
6	2	Research in Librarianship	A. A. L. S. Committee on Research	Oct. 1957
6	3	Building Library Resources Through Cooperation	Ralph T. Esterquest	Jan. 1958
6	4	Legal Aspects of Library Administration	John B. Kaiser	April 1958
V. 7, N. 1		Current Trends in Book Publishing	Frank L. Schick	July 1958
7	2	Aspects of Library Public Relations	Len Arnold	Oct. 1958
7	3	Current Trends in Library Administration	Ernest J. Reece	Jan. 1959
7	4	Current Trends in Bibliography	Roy B. Stokes	April 1959
V. 8, N. 1		Current Trends in Adult Education	C. Walter Stone	July 1959
8	2	Current Trends in Newly Developing Countries	Wilfred J. Plumbe	Oct. 1959
8	3	Photoduplication in Libraries	James E. Skipper	Jan. 1960
8	4	Music Libraries and Librarianship	Vincent Duckles	April 1960
V. 9, N. 1		State Aid to Public Libraries	S. Janice Kee	July 1960
9	2	Current Trends in Theological Libraries	Niels H. Sonne	Oct. 1960
9	3	Current Trends in Bookmobiles	Harold Goldstein	Jan. 1961
9	4	Current Trends in Antiquarian Books	Hellmut Lehmann-Haupt	April 1961
V. 10, N. 1		Future of Library Service: Demographic Aspects and Implications, Part I	Frank L. Schick	July 1961
10	2	Future of Library Service: Demographic Aspects and Implications, Part II	Frank L. Schick	Oct. 1961
10	3	Current Trends in U. S. Periodical Publishing	Helen M. Welch Maurice F. Tauber	Jan. 1962
10	4	Urban University Libraries	Lorena A. Garloch	April 1962
V. 11, N. 1		Library Boards	J. Archer Eggen	July 1962
11	2	Bibliotherapy	Ruth M. Tews	Oct. 1962
11	3	Law Libraries	Bernita J. Davies	Jan. 1963
11	4	Financial Administration of Libraries	Ralph H. Parker Paxton P. Price	April 1963
V. 12, N. 1		Public Library Service to Children	Winifred C. Ladley	July 1963
12	2	Education for Librarianship Abroad in Selected Countries	Harold Lancour J. Clement Harrison	Oct. 1963
12	3	Current Trends in Reference Services	Margaret Knox Goggin	Jan. 1964
12	4	European University Libraries: Current Status and Developments	Robert Vosper	April 1964
V. 13, N. 1		Research Methods in Librarianship	Guy Garrison	July 1964
13	2	State and Local History in Libraries	Clyde Walton	Oct. 1964
13	3	Regional Public Library Systems	Hannis S. Smith	Jan. 1965
13	4	Library Furniture and Furnishings	Frazer G. Poole	April 1965
V. 14, N. 1		Metropolitan Public Library Problems Around the World	H. C. Campbell	July 1965
14	2	Junior College Libraries	Charles L. Trinkner	Oct. 1965
14	3	Library Service to Industry	Katharine G. Harris Eugene B. Jackson	Jan. 1966
14	4	Current Trends in Branch Libraries	Andrew Geddes	April 1966

Complete List of Library Trends Issues in Print

		Title	*Editor*	*Date*
V. 15,	N. 1	Government Publications	Thomas S. Shaw	July 1966
15	2	Collection Development in University Libraries	Jerrold Orne	Oct. 1966
15	3	Bibliography: Current State and Future Trends. Part 1.	Robert B. Downs Frances B. Jenkins	Jan. 1967
15	4	Bibliography: Current State and Future Trends. Part 2		April 1967
V. 16,	N. 1	Cooperative and Centralized Cataloging	Esther J. Piercy Robert L. Talmadge	July 1967
16	2	Library Uses of the New Media of Communication	C. Walter Stone	Oct. 1967
16	3	Abstracting Services	Foster E. Mohrhardt	Jan. 1968
16	4	School Library Services and Administration at the School District Level	Sara K. Srygley	April 1968
V. 17,	N. 1	Group Services in Public Libraries	Grace T. Stevenson	July 1968
17	2	Young Adult Service in the Public Library	Audrey Biel	Oct. 1968
17	3	Development in National Documentation and Information Services	H. C. Campbell	Jan. 1969
17	4	The Changing Nature of the School Library	Mae Graham	April 1969
V. 18,	N. 1	Trends in College Librarianship	H. Vail Deale	July 1969
18	2	University Library Buildings	David C. Weber	Oct. 1969
18	3	Problems of Acquisition for Research Libraries	Rolland E. Stevens	Jan. 1970

Library Trends

Forthcoming numbers are as follows:

July, 1970, *Intellectual Freedom.* Editor: Everett T. Moore, Assistant University Librarian, University of California, Los Angeles.

October, 1970, *State and Federal Legislation for Libraries.* Editor: Alex Ladenson, Acting Librarian, Chicago Public Library.

January, 1971, *Book Storage.* Editor: Mary B. Cassata, Assistant Director for Public Services, Lockwood Memorial Library, State University of Buffalo, Buffalo, New York.

April, 1971, *New Dimensions in Educational Technology for Multi-Media Centers.* Editor: Philip Lewis, President, Instructional Dynamics Incorporated, Chicago, Illinois.